



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 124854**

**TO: Manjunath N Rao**  
**Location: REM-3B81/3C70**  
**Art Unit: 1652**  
**Friday, June 18, 2004**

**Case Serial Number: 10/038723**

**From: Mary Jane Ruhl**  
**Location: Biotech-Chem Library**  
**Remsen 1-A-62**  
**Phone: 571-272-2524**

**maryjane.ruhl@uspto.gov**

### **Search Notes**

Examiner Rao,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl  
Technical Information Specialist  
STIC  
Remsen 1-A-62  
Ext. 22524

**THIS PAGE LEFT BLANK**

124854

From: Chan, Christina  
Sent: Wednesday, June 16, 2004 5:28 PM  
To: Rao, Manjunath N.; STIC-Biotech/ChemLib  
Subject: RE: RUSH sequence search request for 10/038,723

Please rush. Thanks Chris

*Chris Chan*

TC 1600 New Hire Training Coordinator and SPE 1644 & 1642  
(571)-272-0841  
Remsen, 3E89

-----Original Message-----

From: Rao, Manjunath N.  
Sent: Wednesday, June 16, 2004 10:51 AM  
To: Chan, Christina  
Subject: RUSH sequence search request for 10/038,723

Hello Christina,

Please authorize the request below as RUSH. The reason being , this is an amended case and due this bi-week.

Thanks  
-Manjunath

-----  
From: Manjunath N. Rao  
Art Unit 1652, Room 3B81  
Mail Box in Room 3C70  
Phone: 272-0939

Date: 6-16-04

Please search the following as soon as possible for application with serial number  
**10/038,723**

1. SEQ ID NO: 2 and amino acids 1-295 against all commercial amino acid databases,  
issued patents/published applications database and pending application

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE LEFT BLANK**



**database.** Please provide a **print** of all results.

**If you have any questions please call me at the above phone number.**

Thanks

**Manjunath N. Rao**  
**Art Unit 1652, Room 3B81,**  
**Mail Box in Room 3C70,**  
**Remsen Building, USPTO**  
**400, Dulany St.**  
**Alexandria, VA.**  
**Phone: 571-272-0939**

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

**TYPE OF SEARCH:**  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

**VENDOR/COST (where applic.)**  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

THIS PAGE LEFT BLANK

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:43:18 ; Search time 20 Seconds

(without alignments)  
2568.314 Million cell updates/sec

Title: US-10-038-723-2

Sequence: 1 MSFSLALSLGLVCTGLANV.....SKTATASKSTTTRSGMSL 534

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2771	100.0	534	2 A29776	glucan 1,4-alpha-g
2	2742	99.0	640	1 ALASGR	glucan 1,4-alpha-g
3	2742	99.0	640	2 A29166	glucan 1,4-alpha-g
4	2620.5	94.6	639	2 J00607	glucan 1,4-alpha-g
5	2588.5	93.4	639	2 J00479	glucan 1,4-alpha-g
6	1935	69.8	612	2 J01346	glucan 1,4-alpha-g
7	1577	56.9	626	2 S36364	glucan 1,4-alpha-g
8	1577	56.9	626	2 J49625	glucan 1,4-alpha-g
9	1497	54.0	493	2 J06538	glucan 1,4-alpha-g
10	1321	47.7	616	2 S33908	glucan 1,4-alpha-g
11	781.5	28.2	450	2 T39433	glucan 1,4-alpha-g
12	724	26.1	604	2 J09001	glucan 1,4-alpha-g
13	676.5	24.4	519	1 A53549	glucan 1,4-alpha-g
14	597	21.5	549	1 S48474	glucan 1,4-alpha-g
15	522	18.8	778	1 ALBYG	glucan 1,4-alpha-g
16	521	18.8	767	1 J00474	glucan 1,4-alpha-g
17	221.5	8.0	615	2 A64501	glucan 1,4-alpha-g
18	181	6.5	1588	2 A86036	probable adhesin Z
19	181	6.5	1588	2 H91188	probable adhesin E
20	172	6.2	1275	2 G90250	glucan 1,4 alpha g
21	170.5	6.2	1275	2 T33369	hypothetical prote
22	166.5	6.0	107	2 B60754	glucan 1,4-alpha-g
23	161	5.8	1063	2 B66731	hypothetical prote
24	160	5.8	888	2 T46726	secreted acid phos
25	159	5.7	1210	2 A25547	ice nucleation pro
26	159	5.7	12271	2 F90073	hypothetical prote
27	158	5.7	1283	2 T39174	hypothetical serin
28	157	5.7	1258	2 J00188	ice nucleation pro
29	156.5	5.6	2232	2 T34434	hypothetical prote

30	155	5.6	1322	2 S07053	ice nucleation pro
31	154.5	5.6	1200	1 SNPSO	ice nucleation pro
32	153	5.5	1131	2 T41144	hypothetical serin
33	151.5	5.5	1034	2 UC2143	ice nucleation act
34	151.5	5.5	1306	2 S25370	MSB2 protein - yea
35	151.5	5.5	1331	2 A48954	mannan endo-1,4-be
36	150	5.4	1609	2 S25345	probable membrane
37	149	5.4	1609	2 S25345	probable membrane
38	147	5.3	507	2 S64507	hypothetical prote
39	146	5.3	612	2 C90419	probable membrane
40	144.5	5.2	614	2 F66719	glucan 1,4 alpha g
41	144	5.2	948	2 T11678	hypothetical prote
42	142	5.1	1419	2 T30531	hypothetical-like ad
43	141.5	5.1	644	1 T40712	endo-1,4-beta-xy/a
44	141.5	5.1	1441	2 B86807	hypothetical prote
45	140.5	5.1	1567	2 S11672	ice nucleation pro

## ALIGNMENTS

## RESULT 1

A29776  
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) G2 precursor - Aspergillus awamori  
N:Alternate names: glucocamylase G2  
C:Species: Aspergillus awamori  
C:Date: 20-Jun-1989 #sequence\_revision 20-Jun-1989 #text\_change 12-Jun-2003  
C:Accession: A93066; A29776  
R:Numberg, J.H.; Meade, J.H.; Cole, G.; Lawyer, F.C.; McCabe, P.; Schweickart, V.; Tal,  
Mol. Cell. Biol. 4, 2306-2315, 1984  
A:Title: Molecular cloning and characterization of the glucocamylase gene of Aspergillus  
A:Reference number: A93066; MUID:8508534; PMID:6440004  
A:Accession: A93066  
A:Molecule type: DNA  
A:Residues: 1-534 <NDN>  
A:Cross-references: GB:K02465; NID:9454405; PIDN:ABS9297.1; PID:9166506  
R:Numberg, J.H.; Meade, J.H.; Cole, G.; Lawyer, F.C.; McCabe, P.; Schweickart, V.; Tal,  
submitted to Genbank, February 1985  
A:Reference number: A94514  
A:Contents: annotation; revisions to the DNA sequence and coding regions for G2 form  
C:Comment: The DNA sequence was obtained from Genbank, release 55.0.  
C:Superfamily: Glucan 1,4-alpha-glucosidase, Yeast type; glucan 1,4-alpha-glucosidase ho  
C:Keywords: alternative splicing; extracellular protein; glycoprotein; glycosidase; hydr  
F:23-447/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 100.0%; Score 2771; DB 2; Length 534;

Best Local Similarity 100.0%; Pred. No. 5.9e-179;

Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSFSLALSLGLVCTGLANVSKRATLDSMTSNEATVARTALINNIGDGMVSGADSGI	60
DB	1	MSFSLALSLGLVCTGLANVSKRATLDSMTSNEATVARTALINNIGDGMVSGADSGI	60
QY	61	VVASPTNDPDPYFTWPRDGLVLTIVDLPFRNGPTSLSTENTYISQAIVQGISNSPG	120
DB	61	VVASPTNDPDPYFTWPRDGLVLTIVDLPFRNGPTSLSTENTYISQAIVQGISNSPG	120
QY	121	DLSSGAGIGEPKFNVDERTAYTGSWGRPORDPALRATAMIGFQWLNDNGYSTATDIW	180
DB	121	DLSSGAGIGEPKFNVDERTAYTGSWGRPORDPALRATAMIGFQWLNDNGYSTATDIW	180
QY	121	DLSSGAGIGEPKFNVDERTAYTGSWGRPORDPALRATAMIGFQWLNDNGYSTATDIW	180
DB	121	DLSSGAGIGEPKFNVDERTAYTGSWGRPORDPALRATAMIGFQWLNDNGYSTATDIW	180
QY	181	PLVRNDLSYVAQYNQGTGYDLMEEVNGSSFTTIAVQHPALVEGSAFATAVSSSCWCSCQ	240
DB	181	PLVRNDLSYVAQYNQGTGYDLMEEVNGSSFTTIAVQHPALVEGSAFATAVSSSCWCSCQ	240
QY	241	APETICVQSFMTSFTLANPDSRSRGDANTLLGSHITPPPEAACDSTQPCSPRALA	300
DB	241	APETICVQSFMTSFTLANPDSRSRGDANTLLGSHITPPPEAACDSTQPCSPRALA	300
QY	301	NHKEVDSFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPFCTLAABQLYDALYQMD	360
DB	301	NHKEVDSFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPFCTLAABQLYDALYQMD	360

Qy 361 KQSLVETDVSLEDFKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420  
Db 361 KQSLVETDVSLEDFKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420  
Qy 421 SMSEQYDKSDGEQLSARDLTWVSYAALLTANNRRNSVVPASMGETSASSVPGTCAATSAIG 480  
Db 421 SMSEQYDKSDGEQLSARDLTWVSYAALLTANNRRNSVVPASMGETSASSVPGTCAATSAIG 480  
Qy 481 TYSSTVTSWPSIVATGCTTTATPTGSGSVTSSTKTTATASKTSTTTRSGMSL 534  
Db 481 TYSSTVTSWPSIVATGCTTTATPTGSGSVTSSTKTTATASKTSTTTRSGMSL 534

## RESULT 2

ALASGR

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - *Aspergillus niger*

N/Alternate names: 1,4-alpha-D-glucan glucosylhydrolase; glucosylase

C/Species: *Aspergillus niger*

C/Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 02-Jun-2003

C/Accession: A90986; A91161; A05287; A22149; A25402

R/Boel, E.; Hansen, M.T.; Hjort, I.; Hoegh, I.; Fill, N.P.

EMBO J. 3, 1581-1585, 1984

A/Title: Two different types of intervening sequences in the glucosylase gene from *Aspe*

A/Reference number: A90986; MUID:84261458; PMID:6204865

A/Accession: A90986

A/Molecule type: DNA

A/Residues: 1-640 &lt;BOE&gt;

A/Cross-references: GB:X00712; GB:K02466; NID:G2342; PIDN:CAA25303.1; PID:G2343

A/Note: the authors translated the codon GAT for residue 317 as Asn

R/Stenson, B.; Larsen, K.; Gunnarsson, A.

Eur. J. Biochem. 154, 497-502, 1986

A/Title: Characterization of a glucosylase G2 from *Aspergillus niger*.

A/Reference number: A91161; MUID:86136085; PMID:3081341

A/Contents: comparison of forms G1 and G2

A/Accession: A91161

A/Molecule type: protein

A/Residues: 25-640 &lt;SVB&gt;

C/Comment: The large molecular form G1 is shown.

C/Comment: Smaller molecular forms of the enzyme, G2, arise by proteolytic cleavage(s) c

tive towards soluble poly- and oligosaccharides.

C/Genetics: 72/1; 167/3; 200/1; 412/3

C/Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha

C/Keywords: alternative splicing; extracellular protein; glycoprotein; glycosidase; hydr

F:1-24/Domain: signal sequence #status predicted &lt;SIG&gt;

F:23-447/Domain: glucan 1,4-alpha-glucosidase homology &lt;AG&gt;

F:25-640/Product: glucan 1,4-alpha-glucosidase G1 #status experimental &lt;GG1&gt;

F:25-538/Product: glucan 1,4-alpha-glucosidase G2, long form #status experimental &lt;GG3&gt;

F:25-536/Product: glucan 1,4-alpha-glucosidase G2, short form #status experimental &lt;GG2&gt;

F:195-419/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:465-467,468,477,483,484,492,508,510,513,515,522,525,527,529,530,532/Binding site: car

F:476,486,488,496,499,500,501,502,504,506,512,514,517,518,520,524,526,528,531,534,535/Bl

Query Match 99.0%; Score 2742; DB 1; Length 640;  
Best Local Similarity 99.6%; Pred. No. 6.8e-177;

Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSFRSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNIGADGAWVSGADSGI 60  
Db 1 MSFRSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNIGADGAWVSGADSGI 60  
Qy 61 VVASPSTDNPDYFTWTRDSGLVLTIVDLFRNGDTSLSLTINYSIAQAIIVGQISNPSG 120  
Db 61 VVASPSTDNPDYFTWTRDSGLVLTIVDLFRNGDTSLSLTINYSIAQAIIVGQISNPSG 120  
Qy 121 DLSGAGLGPKNVDETATYTGSGRPQRDPALRATAMIGFGOMLLDNGYTSTADTIW 180  
Db 121 DLSGAGLGPKNVDETATYTGSGRPQRDPALRATAMIGFGOMLLDNGYTSTADTIW 180  
Qy 121 DLSGAGLGPKNVDETATYTGSGRPQRDPALRATAMIGFGOMLLDNGYTSTADTIW 180  
Db 121 DLSGAGLGPKNVDETATYTGSGRPQRDPALRATAMIGFGOMLLDNGYTSTADTIW 180  
Qy 181 PIVRNDLSYVAQVWNTGYDLMEEVNGSSPFTIYVQHRALVEGSAFATAVGSSSCSWCDSQ 240  
Db 181 PIVRNDLSYVAQVWNTGYDLMEEVNGSSPFTIYVQHRALVEGSAFATAVGSSSCSWCDSQ 240

Qy 241 APEILCYLQSFMTGSLFANFDSRSGKDNILLGSIHTFDPAAACDSTFQPCSPALA 300  
Db 241 APEILCYLQSFMTGSLFANFDSRSGKDNILLGSIHTFDPAAACDSTFQPCSPALA 300  
Qy 301 NHEEVDSFRSITLNDGLSDSEAVAVGRYPEDTYNNGNWFCTTAAAEQVLDALYQMD 360  
Db 301 NHEEVDSFRSITLNDGLSDSEAVAVGRYPEDTYNNGNWFCTTAAAEQVLDALYQMD 360  
Qy 361 KQSLVETDVSLEDFKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420  
Db 361 KQSLVETDVSLEDFKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420  
Qy 421 SMSEQYDKSDGEQLSARDLTWVSYAALLTANNRRNSVVPASMGETSASSVPGTCAATSAIG 480  
Db 421 SMSEQYDKSDGEQLSARDLTWVSYAALLTANNRRNSVVPASMGETSASSVPGTCAATSAIG 480  
Qy 481 TYSSTVTSWPSIVATGCTTTATPTGSGSVTSSTKTTATASKTSTTTRSGMSL 530  
Db 481 TYSSTVTSWPSIVATGCTTTATPTGSGSVTSSTKTTATASKTSTTTRSGMSL 530

## RESULT 3

A29166

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - *Aspergillus awamori*

N/Alternate names: 1,4-alpha-D-glucan glucosylhydrolase precursor; glucosylase precursor

C/Species: *Aspergillus awamori*

C/Date: 08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change 02-Jun-2003

C/Accession: A29166

R/Nunberg, J.H.; Meade, J.H.; Cole, G.; Lawyer, F.C.; McCabe, P.; Schweickart, V.; Tal, F

Mol. Cell. Biol. 4, 2306-2315, 1984

A/Title: Molecular cloning and characterization of the glucosylase gene of *Aspergillus*

A/Reference number: A93066; MUID:85085934; PMID:6440004

A/Accession: A29166

A/Molecule type: DNA

A/Residues: 1-640 &lt;NUN&gt;

A/Cross-references: GB:K02465; NID:9454405; PIDN:AAB59296.1; PID:9166505

A/Comment: See also PIR:ALASGR.

C/Genetics:

A/Intons: 72/1; 167/3; 200/1; 398/3

C/Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha

C/Keywords: alternative splicing; extracellular protein; glycoprotein; glycosidase; hydr

F:1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F:19-640/Product: glucan 1,4-alpha-glucosidase homology &lt;AG&gt;

F:23-447/Domain: glucan 1,4-alpha-glucosidase homology &lt;AG&gt;

Query Match 99.0%; Score 2742; DB 2; Length 640;  
Best Local Similarity 99.6%; Pred. No. 6.8e-177;

Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSFRSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNIGADGAWVSGADSGI 60  
Db 1 MSFRSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNIGADGAWVSGADSGI 60  
Qy 61 VVASPSTDNPDYFTWTRDSGLVLTIVDLFRNGDTSLSLTINYSIAQAIIVGQISNPSG 120  
Db 61 VVASPSTDNPDYFTWTRDSGLVLTIVDLFRNGDTSLSLTINYSIAQAIIVGQISNPSG 120  
Qy 121 DLSGAGLGPKNVDETATYTGSGRPQRDPALRATAMIGFGOMLLDNGYTSTADTIW 180  
Db 121 DLSGAGLGPKNVDETATYTGSGRPQRDPALRATAMIGFGOMLLDNGYTSTADTIW 180  
Qy 181 PIVRNDLSYVAQVWNTGYDLMEEVNGSSPFTIYVQHRALVEGSAFATAVGSSSCSWCDSQ 240  
Db 181 PIVRNDLSYVAQVWNTGYDLMEEVNGSSPFTIYVQHRALVEGSAFATAVGSSSCSWCDSQ 240  
Qy 241 APEILCYLQSFMTGSLFANFDSRSGKDNILLGSIHTFDPAAACDSTFQPCSPALA 300  
Db 241 APEILCYLQSFMTGSLFANFDSRSGKDNILLGSIHTFDPAAACDSTFQPCSPALA 300  
Qy 301 NHEEVDSFRSITLNDGLSDSEAVAVGRYPEDTYNNGNWFCTTAAAEQVLDALYQMD 360  
Db 301 NHEEVDSFRSITLNDGLSDSEAVAVGRYPEDTYNNGNWFCTTAAAEQVLDALYQMD 360

QY 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTYSIVDAVKTFFADGFSIVETHAASNG 420  
 DB 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTYSIVDAVKTFFADGFSIVETHAASNG 420  
 QY 421 SMSSEYDKSDGQSLARLWTSYALTLANNRRNSVVPASGSETSSASVPGCAATSAIG 480  
 DB 421 SMSSEYDKSDGQSLARLWTSYALTLANNRRNSVVPASGSETSSASVPGCAATSAIG 480  
 QY 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTATASKSTTTSS 530  
 DB 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTATASKSTTTSS 530

RESULT 4  
 J00607  
 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus sp.  
 N:Alternate names: glucosylase  
 C:Species: Aspergillus sp.  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 02-Jun-2003  
 C:Accession: J00607  
 R:Shibuya, I.; Gomi, K.; Iimura, Y.; Takahashi, K.; Tamura, G.; Hara, S.  
 Agric. Biol. Chem. 54, 1905-1914, 1990  
 A:Title: Molecular cloning of the glucanase gene of Aspergillus shirousami and its ex  
 A:Reference number: J00607; MUID:91182400; PMID:1368603  
 A:Accession: J00607  
 A:Molecule type: DNA  
 A:Residues: 1-639 <SHI>  
 A:Experimental source: strain RIB 2504  
 C:Comment: This enzyme catalyzes the release of glucose from the non-reducing ends of st  
 C:Genetics:  
 A:Interons: 72/1; 166/3; 199/1; 411/3  
 C:Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha  
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:23-446/Domain: signal sequence #status predicted <SIG>  
 F:25-639/Product: glucan 1,4-alpha-glucosidase homology <GAG>  
 F:25-639/Product: glucanase #status predicted <MAT>

Query Match 94.6%; Score 2620.5; DB 2; Length 639;  
 Best Local Similarity 94.3%; Pred. No. 1e-168; Indels 1; Gaps 1;  
 Matches 500; Conservative 16; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAVSGADSGI 60  
 DB 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAVSGADSGI 60  
 QY 61 VVASPTNDPFIYTWTRDGLVLTVDLFRNGDTSLSTIENTISAOAIYQGISNPSG 120  
 DB 61 VVASPTNDPFIYTWTRDGLVLTVDLFRNGDTSLSTIENTISAOAIYQGISNPSG 120  
 QY 121 DLSGAGGEPKENVDETAIYGSWGRPORDPALATAMIGGQWLLNGYTSATDIY 180  
 DB 121 DLSGAGGEPKENVDETAIYGSWGRPORDPALATAMIGGQWLLNGYTSATDIY 180  
 QY 121 DLSGAGGEPKENVDETAIYGSWGRPORDPALATAMIGGQWLLNGYTSATDIY 180  
 DB 121 DLSGAGGEPKENVDETAIYGSWGRPORDPALATAMIGGQWLLNGYTSATDIY 180  
 QY 181 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFITIAQHRALVEGSAFATAVSGSCWCSQ 240  
 DB 181 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFITIAQHRALVEGSAFATAVSGSCWCSQ 240  
 QY 180 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFITIAQHRALVEGSAFATAVSGSCWCSQ 239  
 DB 180 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFITIAQHRALVEGSAFATAVSGSCWCSQ 239  
 QY 241 APELLICYQSFWTGSFIILANPDSRSKGADANTLLGSHITPDEAACDDSTPOCSPRALA 300  
 DB 241 APELLICYQSFWTGSFIILANPDSRSKGADANTLLGSHITPDEAACDDSTPOCSPRALA 300  
 QY 240 APELLICYQSFWTGSFIILANPDSRSKGADANTLLGSHITPDEAACDDSTPOCSPRALA 299  
 DB 240 APELLICYQSFWTGSFIILANPDSRSKGADANTLLGSHITPDEAACDDSTPOCSPRALA 299  
 QY 301 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 360  
 DB 301 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 360  
 QY 300 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 359  
 DB 300 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 359  
 QY 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTYSIVDAVKTFFADGFSIVETHAASNG 420  
 DB 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTYSIVDAVKTFFADGFSIVETHAASNG 420  
 QY 421 SMSSEYDKSDGQSLARLWTSYALTLANNRRNSVVPASGSETSSASVPGCAATSAIG 480  
 DB 421 SMSSEYDKSDGQSLARLWTSYALTLANNRRNSVVPASGSETSSASVPGCAATSAIG 480  
 QY 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTATASKSTTTSS 530  
 DB 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTATASKSTTTSS 530

QY 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTATASKSTTTSS 530  
 DB 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTATASKSTTTSS 530

RESULT 5  
 J00479  
 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus awamori  
 N:Alternate names: glucosylase I  
 C:Species: Aspergillus awamori  
 C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 02-Jun-2003  
 C:Accession: J00479  
 R:Hayashida, S.; Kuroda, K.; Ohta, K.; Kuhara, S.; Fukuda, K.; Sakaki, Y.  
 Agric. Biol. Chem. 53, 923-929, 1989  
 A:Title: Molecular cloning of the glucanase I gene of Aspergillus awamori var. kawach  
 A:Reference number: J00479  
 A:Accession: J00479  
 A:Molecule type: DNA  
 A:Residues: 1-639 <HAY>  
 A:Experimental source: var. kawachi  
 C:Genetics:  
 A:Gene: GAI  
 A:Interons: 72/1; 166/3; 199/1; 411/3  
 C:Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha  
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:23-446/Domain: signal sequence #status predicted <SIG>  
 F:25-639/Product: glucan 1,4-alpha-glucosidase #status predicted <MAT>  
 F:494-538/Region: raw-starch-affinity region

Query Match 93.4%; Score 2588.5; DB 2; Length 639;  
 Best Local Similarity 93.8%; Pred. No. 1.5e-166; Indels 1; Gaps 1;  
 Matches 497; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

QY 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAVSGADSGI 60  
 DB 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAVSGADSGI 60  
 QY 61 VVASPTNDPFIYTWTRDGLVLTVDLFRNGDTSLSTIENTISAOAIYQGISNPSG 120  
 DB 61 VVASPTNDPFIYTWTRDGLVLTVDLFRNGDTSLSTIENTISAOAIYQGISNPSG 120  
 QY 121 DLSGAGGEPKENVDETAIYGSWGRPORDPALATAMIGGQWLLNGYTSATDIY 180  
 DB 121 DLSGAGGEPKENVDETAIYGSWGRPORDPALATAMIGGQWLLNGYTSATDIY 180  
 QY 121 DLSGAGGEPKENVDETAIYGSWGRPORDPALATAMIGGQWLLNGYTSATDIY 180  
 DB 121 DLSGAGGEPKENVDETAIYGSWGRPORDPALATAMIGGQWLLNGYTSATDIY 180  
 QY 181 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFITIAQHRALVEGSAFATAVSGSCWCSQ 240  
 DB 181 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFITIAQHRALVEGSAFATAVSGSCWCSQ 240  
 QY 180 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFITIAQHRALVEGSAFATAVSGSCWCSQ 239  
 DB 180 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFITIAQHRALVEGSAFATAVSGSCWCSQ 239  
 QY 241 APELLICYQSFWTGSFIILANPDSRSKGADANTLLGSHITPDEAACDDSTPOCSPRALA 300  
 DB 241 APELLICYQSFWTGSFIILANPDSRSKGADANTLLGSHITPDEAACDDSTPOCSPRALA 300  
 QY 240 APELLICYQSFWTGSFIILANPDSRSKGADANTLLGSHITPDEAACDDSTPOCSPRALA 299  
 DB 240 APELLICYQSFWTGSFIILANPDSRSKGADANTLLGSHITPDEAACDDSTPOCSPRALA 299  
 QY 301 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 360  
 DB 301 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 360  
 QY 300 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 359  
 DB 300 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 359  
 QY 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTYSIVDAVKTFFADGFSIVETHAASNG 420  
 DB 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTYSIVDAVKTFFADGFSIVETHAASNG 420  
 QY 421 SMSSEYDKSDGQSLARLWTSYALTLANNRRNSVVPASGSETSSASVPGCAATSAIG 480  
 DB 421 SMSSEYDKSDGQSLARLWTSYALTLANNRRNSVVPASGSETSSASVPGCAATSAIG 480  
 QY 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTATASKSTTTSS 530  
 DB 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTATASKSTTTSS 530

RESULT 6  
 J01346

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - *Aspergillus oryzae*  
N/Alternate names: Glucoamylase  
C/Species: *Aspergillus oryzae*  
C/Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 02-Jun-2003  
C/Accession: J01346  
R/Hata, Y.; Tsuchiya, K.; Kitamoto, K.; Gomi, K.; Kumagai, C.; Tamura, G.; Hara, S.  
Gene 108, 145-150, 1991  
A/Title: Nucleotide sequence and expression of the glucoamylase-encoding gene (glua) frc  
A/Reference number: J01346; MUID:92104497; PMID:1761224  
A/Accession: J01346  
A/Molecule type: DNA  
A/Residues: 1-612 <NAT>  
A/Cross-references: GB:D10698; DDBJ:D01108; NID:g1160312; PIDN:BA01540.1; PID:g1160313  
A/Note: the authors translated the codon TTT for residue 213 as Tyr  
C/Comment: This enzyme hydrolyzes starch to glucose.  
C/Genetics:  
A/Gene: glua  
A/Intons: 75/1; 169/3; 202/1; 414/3  
C/Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha  
C/Keywords: glycosidase; hydrolase; polysaccharide degradation  
F:26-449/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 69.8%; Score 1935; DB 2; Length 612;  
Best Local Similarity 70.3%; Pred. No. 1,4e-122;  
Matches 367; Conservative 59; Mismatches 90; Indels 6; Gaps 4;

Qy 1 MSFRLSLALSGVCTGLA--NVISKRAITDMSLSEATVARTAILNINIGADGAWVGADGS 58  
Db 2 VSPSSCLRALALGSSVLAQPVLRQATGDLDTLSTEANFSQAILNINIGADGSAQAQASP 61  
Qy 59 GIVVASPSTNDPFFYTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAVQGISNP 118  
Db 62 GVVIASPSKSDPDFFYTWTRDSGLVMTKTVLDFRGADADLLPIIEEFISSQARIQGISNP 121  
Qy 119 SGLSSGAGGEPENDEFAATYNGSKRPQRDGPALATAMIGRCOMLNDNGYSTATDI 178  
Db 122 SGALSSG-GGEPENDEFAATYNGSKRPQRDGPALATAMISFGEWLVENSHTSIATDL 180  
Qy 179 VWPVLRDLTVYAQYMNQGYDLMEEVNCGSFFTIAYOHRALVEGSAFATAVSSGCMCD 238  
Db 181 VWPVLRDLTVYAQYMNQGYDLMEEVNCGSFFTIAYOHRALVEGSAFATAVSSGCMCD 240  
Qy 239 SQAPEILCYLQSFMTGSPFIILNFDSSRSKQDANTLGSIHTEDEAACDDSTFQPCSPRA 298  
Db 241 SQAPQVQCYLQSFMTGSPFIILNFDSSRSKQDANTLGSIHTEDEAACDDSTFQPCSPRA 300  
Qy 299 LANHKEVVDSPFRSYLTINDGLSDEAVAVGRYPEDTYNNGPWFLCTILAAAEQLYDALYQ 358  
Db 301 LANHKEVVDSPFRSYLTINDGLSDEAVAVGRYPEDTYNNGPWFLCTILAAAEQLYDALYQ 360  
Qy 359 WDKQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIYDAVKTADGFSIVETHAAS 418  
Db 361 WDKIGSLAITDVSLDFPKALYSMAITGYSSSTYVYKDIYSAVAYADYQVQIYQIYTAAS 420  
Qy 419 NGWSEQYKSDGEQLSARDLTWMSYALLTANNRRNSVPAWGETSASSVPGTCAATSA 478  
Db 421 TGSMAEQYTKDQGSQTSARDLTWMSYALLTANNRRNSVPAWGETSASSVPGTCAATSA 480  
Qy 479 IGTYSSTVTWSPSIVATGCT--TTTATPTGSGSVTSTSKTT 518  
Db 481 SGTYSSTVTWSPSIVATGCT--TTTATPTGSGSVTSTSKTT 521

## RESULT 7

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - *Neurospora crassa*  
N/Alternate names: Glucoamylase; glycoamylase  
C/Species: *Neurospora crassa*  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 02-Jun-2003  
C/Accession: S36364; S13710; S13711; S25539  
R/Stone, P.J.; Makof, A.J.; Parish, J.H.; Radford, A.  
Curr. Genet. 24, 205-211, 1993  
A/Title: Cloning and sequence analysis of the glucoamylase gene of *Neurospora crassa*.

A/Reference number: S36364; MUID:94037144; PMID:8221928  
A/Accession: S36364  
A/Molecule type: DNA  
A/Residues: 1-626 <STO>  
A/Cross-references: EMBL:X67291  
R/Koh-luar, S.I.; Parish, J.H.; Bleasby, A.J.; Pappin, D.J.C.; Ainley, K.; Johansen, F.E.  
Enzyme Microb. Technol. 11, 692-695, 1989  
A/Title: Exported proteins of *Neurospora crassa*: 1-glucoamylase.  
A/Reference number: S13710  
A/Accession: S13710  
A/Molecule type: protein  
A/Residues: 36-60, 'X', 62, 'X', 64-65 <KOH>  
R/Koh-luar, S.I.; Parish, J.H.; Bleasby, A.J.; Pappin, D.J.C.; Ainley, K.; Johansen, F.E.  
submitted to the Protein Sequence Database, January 1990  
A/Description: Exported proteins of *Neurospora crassa* 1: - glucoamylase.  
A/Reference number: S13711  
A/Accession: S13711  
A/Molecule type: protein  
A/Residues: 36-60, 'X', 62, 'X', 64-65 <KOH>  
C/Genetics:  
A/Gene: glua-1  
A/Intons: 82/2  
C/Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha  
C/Keywords: glycosidase; hydrolase; polysaccharide degradation  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-35/Domain: propeptide #status predicted <PRO>  
F:33-458/Domain: glucan 1,4-alpha-glucosidase homology <GAG>  
F:36-626/Product: glucan 1,4-alpha-glucosidase #status experimental <MAT>

Query Match 56.9%; Score 1577; DB 2; Length 626;  
Best Local Similarity 57.6%; Pred. No. 1.9e-98;  
Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

Qy 2 SPSRLALSGVCTGLANVSKRAITDMSLSEATVARTAILNINIGADGAWVGADGIV 61  
Db 13 AFDVAGLDPDLPEKRSIDILKA-SVDSYIQETPTIAQKMLCNIGASGRASGAASGV 71  
Qy 62 VASPTNDPFFYTWTRDSGLVLTVDLFRNG-DTSLSTIENYISAQAVQGISNP 120  
Db 72 VASPSKSSPDYTWTRDALVTKLIYDEFTNNTQITQIYAAQAQKLGVSNSPG 131  
Qy 121 DLSGAGLGPKNVDETATYNGSKRPQRDGPALATAMIGRCOMLNDNGYSTATDIY 180  
Db 132 SLNSGAGLGPKNVMDLQCFTGAMGRPQRDGPALATAMIGRCOMLNDNGYSTATDIY 191  
Qy 181 PLVNRDLTVYAQYMNQGYDLMEEVNCGSFFTIAYOHRALVEGSAFATAVSSGCMCD 240  
Db 192 PIVNRDLTVYAQYMNQGYDLMEEVNCGSFFTIAYOHRALVEGSAFATAVSSGCMCD 251  
Qy 241 APEILCYLQSFMTGSPFIILNFDSSRSKQDANTLGSIHTEDEAACDDSTFQPCSPRAL 299  
Db 252 APEILCYLQSFMTGSPFIILNFDSSRSKQDANTLGSIHTEDEAACDDSTFQPCSPRAL 311  
Qy 300 ANHKEVVDSPFRSYLTINDGLSDEAVAVGRYPEDTYNNGPWFLCTILAAAEQLYDALYQ 359  
Db 312 ANHKEVVDSPFRSYLTINDGLSDEAVAVGRYPEDTYNNGPWFLCTILAAAEQLYDALYQ 370  
Qy 360 DKGSLLEVTDVSLDFPKALYSDAATGYSSSSSTYSIYDAVKTADGFSIVETHAAS 419  
Db 371 KKGSLLEVTDVSLDFPKALYSDAATGYSSSSSTYSIYDAVKTADGFSIVETHAAS 430  
Qy 420 GSWSEQYKSDGEQLSARDLTWMSYALLTANNRRNSVPAWGETSASSVPGTCAATSAI 479  
Db 431 GSWSEQYKSDGEQLSARDLTWMSYALLTANNRRNSVPAWGETSASSVPGTCAATSAI 490  
Qy 480 GTYSSTVTWSPSIVATGCTTT-ATPTGSGS-----VTSSTKTATAGT 524  
Db 491 GSYATATATSPANLTPASTTVPPTQTCGAADHEVLVTENKVTTSYQGT 541

## RESULT 8

T49625  
glucan 1,4-alpha-glucosidase [imported] - *Neurospora crassa*

N:Alternate names: protein B5022.70  
 C:Species: Neurospora crassa  
 C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2003  
 C/Accession: T49625  
 R/Schulte, U.; Algn. V.; Hohenstein, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000  
 A/Reference number: 225022  
 A/Accession: T49625  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-626 <SCH>  
 A/Cross-references: EMBL:AL359932; GSPDB:GN00116; NCSP:B5022.70  
 A/Experimental source: BAC clone B5022, strain OR74A  
 C/Genetics:  
 A/Gene: NCSP:B5022.70  
 A/Map position: 6  
 A/Insertion: 82/1  
 C/Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha

Query Match 56.9%; Score 1577; DB 2; Length 626;  
 Best Local Similarity 57.6%; Pred. No. 1.9e-98;  
 Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

2 SFRSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWTSGADSGIV 61  
 13 ARQAVLGPRDPIHEKRHSIIKR-SVDSYIQTETPIAQNKLICNGASGCRASGAASGV 71  
 62 VASPTDNDFYTTWRDGLVLTVDLFRNG-DTSLSTIENTYISQAIVQGISNPSG 120  
 72 VASPSKSPDYWTYTRDALVTKLIVDEFNDYNTLONTIQAVALAQAQKQGSNPSG 131  
 121 DLSAGAGEPKENDETAYTSGMGRPOBDGPALRATAMIGFQWILNDGYSTADTV 180  
 132 SLNSNGGLSEEPFMDVLDQFPGAMGPRQDGPPLRALIGYKMLVNSGVADTKSLIT 191  
 181 PLVRNDLSVAQYNNQGYDLMEVNGSSFTIAVOHRLVGSAPATVGSSCSWCDSQ 240  
 192 PIVKNDLAVTAQYNNNTGDLMEVNGSSFTIAASHRLVGSAPATVGSSCSWCDSQ 251  
 241 APELICQYQSWTGS-FILANFDSRSRSGKANTLGSITHTPEPAACDDSTFQPCSPRAL 299  
 252 APQILCFQOOSFMSNGYIISFNFNYSRSGKDINSVLTSIHNPDAAGCDVNTFQPCSDRAL 311  
 300 ANHKVNSFRSITLNLGSLDSEAVAVGRYPEDTYNNPWFCTLLAAEQLYDALY 359  
 312 ANHKVNVDSMR-FWGNVSGRTAGKAAAVGRYADYVYNNPWLTLAAEQLYDALY 370  
 360 DKQGLLEVTVSLDFPKALYSDAATGYSSSSSTYSIYDAVKTADGVSIIVETHAASN 419  
 371 KKQGSITTTSTSLAFKDLVPSTGYSSSSSTYTAIINAVTTADGVVDIYAOYTTS 430  
 420 GSMSEQYDKSDGEOLSARDLWTSYALLTANNRRNSVVPASGGETSASSVPCTCATSAI 479  
 431 GSLAQFQDKSDGAPSLATHTLWTSYASFLSAAARRAGIYVPSGGAASANSLPSCSASTYA 490  
 480 GTYSSVTYTSWPSIYATGTTT-ATPTGSGS-----VTSTKTRATASKT 524  
 491 GSATATATISFPANLPASTYVTPPTQTGCADHEVLTFENKVTTSYQOT 541

RESULT 9  
 JG6538  
 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus oryzae  
 C/Species: Aspergillus oryzae  
 C/Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 12-Jun-2003  
 C/Accession: JG6538  
 R/Hata, Y.; Ishida, H.; Ichikawa, E.; Kawato, A.; Suginami, K.; Imayasu, S.  
 Gene 207, 127-134, 1998  
 A/Title: Nucleotide sequence of an alternative glucocamylase-encoding gene (glab) express  
 A/Reference number: JG6538; MUID:98172744; PMID:9511753  
 A/Accession: JG6538  
 A/Molecule type: protein  
 A/Residues: 43-53;71-82;136-145;180-187;327-336;393-404 <HAZ>

A/Experimental source: O-1013  
 A/Accession: JG6538  
 A/Molecule type: mRNA  
 A/Residues: 1-493 <HAZ>  
 A/Cross-references: DDBJ:AB007825  
 C/Comment: This enzyme is used in industrial processing. It releases glucose from the no  
 C/Genetics:  
 A/Insertion: 75/1; 344/2  
 C/Superfamily: glucan 1,4-alpha-glucosidase, yeast type; glucan 1,4-alpha-glucosidase ho  
 C/Keywords: glycoprotein; glycosidase; hydrolase  
 F/1-26/Domain: signal sequence #status predicted <SIG>  
 F/26-450/Domain: glucan 1,4-alpha-glucosidase homology <GAG>  
 F/27-493/Product: glucan 1,4-alpha-glucosidase #status experimental <MAT>  
 F/139,198,255,369,384,457/Binding site: carbohydrate (am) (covalent) #status predicted

Query Match 54.0%; Score 1497; DB 2; Length 493;  
 Best Local Similarity 56.9%; Pred. No. 3.3e-93;  
 Matches 277; Conservative 80; Mismatches 128; Indels 2; Gaps 2;

5 SLTALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWTSGADSGIV 64  
 8 SLNATAGVAPHPSPFIRKQSDNLNFIQAQPIAQTGILNNGADGKLVGGAAAGIVAS 67  
 65 PSTDNPDIYTTWRDGLVLTVDLFRNG-DTSLSTIENTYISQAIVQGISNPSG 124  
 68 PSKSNPDYTTWRDGLVLTVDLFRNG-DTSLSTIENTYISQAIVQGISNPSG 127  
 125 GAGLGEPRVNDDETAYTSGMGRPOBDGPALRATAMIGFQWILNDGYSTADTV 184  
 128 GSGLAEPKFFYVYISQFTDSWGRPQDGPALRASALIAVGNLSISDKQSVVXANIMPV 187  
 185 NDLSVAQYNNQGYDLMEVNGSSFTIAVOHRLVGSAPATVGSSCSWCDSQ 244  
 188 NDLSVAGYNNQGYDLMEVNGSSFTIAVOHRLVGSAPATVGSSCSWCDSQ 246  
 245 LCYLQSPFTGSFILANFDS-RSGKANTLGSITHTPEPAACDDSTFQPCSPRAL 303  
 247 LCHLQDFMNGSVNLNLTNGRSGIDTSLGSIHTTPPAAACDDSTFQPCSPRAL 306  
 304 EYVDSFRSITLNLGSLDSEAVAVGRYPEDTYNNPWFCTLLAAEQLYDALY 363  
 307 LVVDSFRSITLNLGSLDSEAVAVGRYPEDTYNNPWFCTLLAAEQLYDALY 366  
 364 SLEVTVDLDFPKALYSDAATGYSSSSSTYSIYDAVKTADGVSIIVETHAASN 423  
 367 QNVATETSLPFFKDLSPSTGYSSSSSTYTAIINAVTTADGVVDIYAOYTTS 426  
 424 EYVDSDEQYDKSDGEOLSARDLWTSYALLTANNRRNSVVPASGGETSASSVPCTCATSAI 483  
 427 EYVDSDEQYDKSDGEOLSARDLWTSYALLTANNRRNSVVPASGGETSASSVPCTCATSAI 486  
 484 SVTVTSW 490  
 487 TPTVGSW 493

RESULT 10  
 S33908  
 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) P precursor - creosote fungus  
 N:Alternate names: glucocamylase  
 C/Species: Amorphotheca resiniae, Hormoconis resiniae (creosote fungus)  
 C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 02-Jun-2003  
 C/Accession: S33908; S36577; S31423; A60754; S31422  
 R/Vanholo, A.E.I.; Torckell, H.T.; Thues, T.; Aho, S.A.; Fagerstrom, B.R.; Korhola, M.P.  
 Curr. Genet. 24, 38-44, 1993  
 A/Title: Cloning and expression of Hormoconis resiniae glucocamylase P cDNA in Saccharomyces  
 A/Reference number: S33908; MUID:93365035; PMID:8358830  
 A/Accession: S33908  
 A/Molecule type: mRNA  
 A/Residues: 1-616 <VAL1>  
 A/Cross-references: EMBL:X67708; NID:g2785; PIDN:CAA47945.1; PID:g2786  
 A/Accession: S36577  
 A/Molecule type: protein

A;Residues: 30-106;138-151,168-181,217-223,260-294;310-316,321-334,336-342,369-415,423-4  
 R;JoutelJOKI, V.; Torkkeli, T.  
 submitted to the EMBL Data Library, August 1992  
 A;Reference number: S31425  
 A;Accession: S31425  
 A;Molecule type: DNA  
 A;Residues: 1-616 <JOU>  
 A;Cross-references: EMBL:X69143; NID:G2787; PIDN:CAA48243.1; PID:G2788  
 R;Fagerstrom, R.; Valino, A.; Suoranta, K.; Pakula, T.; Kalkkinen, N.; Torkkeli, H.  
 J. Gen. Microbiol. 136, 913-920, 1990  
 A;Title: Comparison of two glucoamylases from *Homocidus resiniae*.  
 A;Reference number: A60754; MUID:90338987; PMID:2116499  
 A;Accession: A60754  
 A;Molecule type: protein  
 A;Residues: 30-53;72-89;145-154;217-223;260-294;423-426,'X',428-436 <FAG>  
 A;Experimental source: strain ATCC 20495  
 C;Genetics:  
 A;Gene: gamp  
 A;Introns: 76/1; 177/3; 347/2  
 C;Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha  
 C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation  
 F;1-29/Domains: signal sequence #status predicted <SIG>  
 F;27-45/Domain: glucan 1,4-alpha-glucosidase homology <GAG>  
 F;30-616/Product: glucan 1,4-alpha-glucosidase #status experimental <MAT>  
 F;99,427/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;200,514,528,587/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 47.7%; Score 1321; DB 2; Length 616;  
 Best Local Similarity 48.6%; Pred. No. 3.2e-81;  
 Matches 257; Conservative 88; Mismatches 150; Indels 34; Gaps 6;

Qy 14 CTGLANVIS-----KRATLDWLSNEATVARTALINNIGADGAWSGADSGIVA 63  
 Db 8 CAGALSLICSLAIAPELTKARDLSSFTASRAIALQCALNNIGPDGSAVPGAGGFVA 67

Qy 64 SPSTDNPDYFTYTRDGLVLTVDLFRNGDTSLSLTENYISAQAIYVGISNPSGD-L 122  
 Db 68 SPSTANPDYFTYTRDGLVLTVDLFRNGDTSLSLTENYISAQAIYVGISNPSGTF 127

Qy 123 SSGAGLSEPKFNVDETAATGSGRPPORDGPALRATAMGFGOMLLNDGTSATDVIWPL 182  
 Db 128 PDGVLSEPKFNVDETAATGSGRPPORDGPALRATAMGFGOMLLNDGTSATDVIWPL 187

Qy 183 VRNDSLVAQYWNQGYDLMEEVNGSSEFTIYVQHRALVEGSAFATAVGSSCSMCDQAP 242  
 Db 188 IANDLSYQYWNQSGDFLMEETVYASSFTIQNGHRALVEGAQALHDLGYVCTGCD-QAP 246

Qy 243 EILCYLQSFYTGSTILANF--DSSRSKGDANTLIGSTHTPEPAACDSTFQPCSPRALA 300  
 Db 247 EVLCLFLOSFMNGKIVSNINNNRGTGDNSTIGALSTFDIDAYCDSPFLQPCQSOLA 306

Qy 301 NHEKVVDSFRSITLNDGLDSEBAVAVGRYPEDTYNNGNPFCLTLLAAEQLYDALYQMD 360  
 Db 307 NFKVLDTFRMLYTNMGIPGCGVAVGRVYAEVDYMGSNPFYLLTTAAEFLDVAQWVK 366

Qy 361 KQGLSEVTVDSLDFFKALYSDAATGYSS--SSSTYSIVDAVVTAFDFSVIETHAAS 418  
 Db 367 ARHVLTVDETSIAFFKDIYEVTVREYKSGNANSPFAQIMAVAYADSVYALIEKPIPS 426

Qy 419 NGMSSEQYDKSDGEQLSARDLTWSYALILTANRRNSVVPASKEVTSASVPGICATSA 478  
 Db 427 NGSLSEQFNRTYGPLSLAIDLTWSSYAAFTTWSORRAGQPSWSSRNALPPPTCSASST 486

Qy 479 IGTSSVTVTWSPIVATAGTTTATPTGSGVSTSTSKTTPASTSTT 527  
 Db 487 PGTY-----TPATAGAPNVVSSCOVSTTFINMTT 517

RESULT 11  
 T39433  
 glucoamylase precursor - fission yeast (*Schizosaccharomyces pombe*)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 12-Jun-2003

C;Accession: T39433  
 R;Xiang, Z.; Aves, S.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volkert, G.  
 submitted to the EMBL Data Library, March 1998  
 A;Reference number: 221854  
 A;Accession: T39433  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-450 <LYN>  
 A;Cross-references: EMBL:AL022305; PIDN:CAA18423.1; GSPDB:GN00067; SPDB:SPBC14C8.05C  
 A;Experimental source: strain 972h-; cosmid c14C8  
 C;Genetics:  
 A;Gene: SPDB:SPBC14C8.05C  
 A;Map position: 2  
 C;Superfamily: glucan 1,4-alpha-glucosidase, yeast type; glucan 1,4-alpha-glucosidase hon

Query Match 28.2%; Score 781.5; DB 2; Length 450;  
 Best Local Similarity 37.6%; Pred. No. 4.4e-45;  
 Matches 179; Conservative 68; Mismatches 166; Indels 63; Gaps 12;

Qy 7 LALSLGVCTG---LANVISKRATLDWLSNEATVARTALINNIGADGAWSGADSGIVA 63  
 Db 8 LLLGVVASLSLSPNKRSEASDEWTDQKGLAMGHMLNIGDSGMHAKDINPGCTIA 67

Qy 64 SPSTDNPDYFTYTRDGLVLTVDLFRNGDTSLSLTENYISAQAIYVGISNPSGDLS 123  
 Db 68 SPSTSDPDYTYQWRDGLVLTVDLFRNGDTSLSLTENYISAQAIYVGISNPSGDFY 127

Qy 124 SGAGLSEPKFNVDETAATGSGRPPORDGPALRATAMGFGOMLLNDGTSATDVIWPL 183  
 Db 128 AG-GLSEPKFNVDETAATGSGRPPORDGPALRATAMGFGOMLLNDGTSATDVIWPL 186

Qy 184 RNDLSVAQYWNQGYDLMEEVNGSSEFTIYVQHRALVEGSAFATAVGSSCSMCDQAP 243  
 Db 187 LADLDYANMTWTEASFDLMEIEKDVHYFTLAVQKRAMQDGTAFKRG-----APD 237

Qy 244 -----ILCYLQSFY-TGSTILANFDS--SRGKQANTLIGSTHT--FDPEAACD 287  
 Db 238 QALYKRTIEPIDLKGEFMDPGVYKGYKGRVDRGLDCLSTLASLYSNEPD----- 291

Qy 288 DSTFQPCSPALANHKEVNSFRSITLNDGLDSEBAVAVGRYPEDTYN-----GNPWF 342  
 Db 292 -----MHLPLTLKQSTMTDRDYPVNGMKO---ANGKRPEDYDVGVSIGNPF 339

Qy 343 LCTLLAAEQLYDALYQMDKQSLSEVTVDSLDFFK--ALYSDAATGYVSSSSSTYSIVDA 400  
 Db 340 ICTSSAEIILYKALAYVDNKGELPELTETYNLHFFKFAEPD-----PYMWSVIRKN 390

Qy 401 VKTFADGFVSIETHAASNGSEQYDKSDGEQLSARDLTWSYALILTANRRNSV 456  
 Db 391 MHTYADNFKLAVAEFHPHNGSMSBQFSDRDGHQGARDLTWSYSSSLTALAYRBEAI 446

RESULT 12  
 JP0001  
 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - *Rhizopus oryzae*  
 N;Alternate names: 1,4-alpha-D-glucan glucanhydrolase; glucoamylase  
 C;Species: *Rhizopus oryzae*  
 C;Date: 28-Dec-1987 #sequence\_revision 23-Aug-1996 #text\_change 16-Jun-2000  
 C;Accession: JP0001  
 R;Ashikari, T.; Nakamura, N.; Tanaka, Y.; Kiyuchi, N.; Shihano, Y.; Tanaka, T.; Amachi, T.  
 Agric. Biol. Chem. 50, 957-964, 1986  
 A;Title: Rhizopus raw-starch-degrading glucoamylase: its cloning and expression in yeast  
 A;Reference number: A90022  
 A;Accession: JP0001  
 A;Molecule type: DNA  
 A;Residues: 1-604 <ASH>  
 A;Cross-references: GB:D00049; PIDN:BA00033.1; PID:G218035  
 A;Experimental source: strain SAM0034  
 A;Note: there are two errors in the published sequence (personal communication): GCT (11)  
 Y to the known amino acid sequence of 11 peptides from glucoamylase, including the amino-  
 R;Tanaka, Y.; Ashikari, T.; Nakamura, N.; Kiyuchi, N.; Shihano, Y.; Amachi, T.; Yoshizumi,  
 Agric. Biol. Chem. 50, 965-969, 1980  
 A;Title: Comparison of amino acid sequences of three glucoamylases and their structure-f



A:Reference number: A90023

A:Contents: annotation; homology; predicted secondary structure

C:Comment: Rhizopus glucoamylase exists in multiple forms, Gluc 1, Gluc 2, and Gluc 3, and degraded raw starch.

C:Genetics:

A:Intron: 51/3, 110/3, 129/1, 436/3

C:Superfamily: Rhizopus glucan 1,4-alpha-glucosidase; glucan 1,4-alpha-glucosidase homol

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-604/Product: glucan 1,4-alpha-glucosidase (Gluc 1) #status predicted <GL1>

F:34-138/Domain: glucan 1,4-alpha-glucosidase (Gluc 2) #status predicted <GL2>

F:116-604/Product: (or 110-604) glucan 1,4-alpha-glucosidase (Gluc 3) #status predicted <GL3>

F:159-604/Product: glucan 1,4-alpha-glucosidase (Gluc 2) #status predicted <GL2>

F:166-592/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 26.1%; Score 724; DB 1; Length 604;

Best Local Similarity 36.8%; Pred. No. 5e-41;

Matches 165; Conservative 71; Mismatches 170; Indels 42; Gaps 10;

25 ATLDSWLSNEATVARTALINNIGADGAWSGADSGIVASPTDNDPYYTTRDGL 84

168 STISSMIKKQEGISRPALMRNINP-----PGSATGFIASLSTAGDYYTATRDALTS 222

85 KTLVDLFR---NGDTSLSSTIENYISAQAVGQISNPSGDLSSGAGLGEPRFNVDETAT 141

223 NVIVVEYNTTSGNKTILNVLDYTFSTKSTSTVCN-----CLGEPRKTPDGSCT 276

142 GSWGRPQRDGPALRATAMIGF-QWLDNGYSTATDIWPLVRNDLSVYAYNQTYD 200

277 GAMGRPQNDGPALRATATFTFLFADSYLTQTKDASVYTGTLKPAIFKDLDDVYVWVMSNGCD 336

201 LMEEVNGSFFFTIANVQHRALVEGSAFATVGS--CSWCDSPAELCTLOSGFTGSRFL 258

337 LMEEVNGVHFYLYLWARKLLGADPAKKNGSTRASTVYSTASTANKISSFWWSNNW 396

259 ANFDS-----SRSGKANTL-----LGSHTFDPEAKCDSTFOPCSPRALNKHVVDS 308

397 IOVSQSVYGVGSKGLDVTLLAANLGSV-----DDGFTFESKILTAVAVEDS 447

309 FRSITVTLNDGLSDSAVAVGRYPEDTY-----YNGNPWFCLTAAAEQLYALYQWMDKG 363

448 FASLVPINKLPSYLGNSIGRYPEPTYNGNGNSQGNMFLAVTGAELVYRAIKEMWIGN 507

364 SLEVTDSLDPEFKALYSDAATG-TYSSSSSTYSIVDAVKTADGFSIVETHAASNGSM 422

508 GTVVSISLPPFKKDDSSATSGKTYVGSIDNNLAQNALAADRFSLTQJLHANNSGL 567

423 SEQYKSDGQELASARDLTWVYALLTAN 450

568 AEEFRTTGLSTGARDLTWVSHSLITAS 595

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

A:Reference number: S14596

A:Accession: S14596

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-271, 'T', 273-519 <HO2>

A:Cross-references: EMBL:X58117

C:Genetics:

A:Gene: GLA1

C:Superfamily: glucan 1,4-alpha-glucosidase, yeast type; glucan 1,4-alpha-glucosidase hc

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:41-506/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 24.4%; Score 676.5; DB 1; Length 519;

Best Local Similarity 36.1%; Pred. No. 6.4e-38;

Matches 175; Conservative 71; Mismatches 176; Indels 63; Gaps 16;

24 RATLDSWLSNEATVARTALINNIGADGAWSGADSGIVASPTDNDPYYTTRDGL 82

42 RDLDETFLDKQKQVSLVYLLQNIAPBEGQFNDGVGVIAASPTSNPDYVQWTRDSAIT 101

83 ---VAKTVDLFRNGDTSLSSTIENYISAQAVGQISNPSG--DLSSGAGLGEPRFNVDE 137

102 FLTVLSELD--NNFTTLAAVEYYINTSYNLQRTSNPSGSPDENHKGLEPRFNTDG 159

138 TAYTSGMRPQRDGPALRATAM-----IGFGQML-----DNGYSTATDIWPLVR 184

160 SATYGAWRPQNDGPALRATAMISRYLNDVNSLNKGLVLTDSGDIINFEST-EDYKAITK 218

185 NDLSTVAYNQTYDLMEEVNGSFFFTIANVQHRALVEGSAFATV-----GSSCSWCD 238

219 PDLEVIVGYMSTGDLMEENQGRHFFTSVQQRAL-----AYAVDIASPDGDPAFYLD 274

239 SQAPEITCYLOSFWTG-----SFLANFD-----SSRSGKANTLGSHTFDPEAACDD 288

275 STASTLESYLSGSDGCFNVTDVNHIVENPDLQQRSGRLSATYIGPLTHTD-IGESSS 333

289 STFOPCSPRALNKHVVDSFRSITYTLNDGLSDSAVAVGRYPEDTY-----YNGNPWF 343

334 TPEVDNERYVLOSYYLLLEDNKRYSVNSAV--SKAALIGRYPEPTYNGNGNSQGNMFL 391

344 CTLAAEQLYDALYQWMDKQSLLEV--DVSILDFKALYSDAAT-----GTYSSSS 391

392 ATAYAAQVPLYLVND-AKSAANDITINKINDDFKXYIVDLSTINSQVSSDSVTIKGS 450

332 STYSSIVDAVKTADGFSIVETHAASNGSSEQYKSDGQELASARDLTWVYALLTANN 451

451 DEFNTVADNLVTFGDSFLQVILIDHINDGSLNEQLNRYGTGYSTSAVSLTWSSGALLBAIR 510

452 RNSV 456

511 LNKV 515

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

A:Residues: 1-503,507,'W',513-514,'TG',516 <YAM>  
 A:Cross-references: EMBL:M16166; NID:G172592; PIDN:AAA5042.1; PID:G172593  
 R:Pardo, J.M.; Ianez, E.; Zalacain, M.; Claros, M.G.; Jimenez, A.  
 FEBS Lett. 239, 179-184, 1988  
 A:Title: Similar short elements in the 5' regions of the STA2 and SGA genes from Sacchar  
 A:Reference number: S27281; MUID:89031230; PMID:3141213  
 A:Accession: S27284  
 A:Molecule type: DNA  
 A:Residues: 1-183,'H',185-190 <PAR>  
 A:Cross-references: EMBL:X13585; NID:G4461; PIDN:CAA32071.1; PID:G4463  
 A:Gene: SGD:SGAL; MIPS:YII099W  
 A:Cross-references: SGD:S0001361; MIPS:YII099W  
 A:Map position: 9L  
 C:Function:  
 A:Description: hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively fr  
 C:Keywords: glucan 1,4-alpha-glucosidase, yeast type; glucan 1,4-alpha-glucosidase ho  
 C:Superfamily: glucan 1,4-alpha-glucosidase, yeast type; glucan 1,4-alpha-glucosidase ho  
 C:Words: glycosidase; hydrolase; polysaccharide degradation; yeast vacuole  
 F:77-536/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 21.5%; Score 597; DB 1; Length 549;  
 Best Local Similarity 30.4%; Pred. No. 1.6e-32;  
 Matches 153; Conservative 89; Mismatches 202; Indels 60; Gaps 14;

Qy 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVSGADSGI 60  
 Db VQLADAVLNGTGYVD-SNGAMDSALBWLQCKKYSIRIFENIGPSAVYPS-ILPGV 113  
 Qy 61 VVASPSTNDPDYFTYTRDGLVLTVDLFRNGDTSLSSTIENYSAQAIIVQGISNPSG 120  
 Db 114 VIAPSPQTHPDYFYQWIRDSALTTINSIVS--HSAGPAIEYLLQYLVNSFHLQRTNN--- 167  
 Qy 121 DLSSGAG-----LGEPRNVDEATYSGRPGQPDGAPALRATAMIGFGWLLDNG--- 170  
 Db 168 TLGAGIGYNTDVTALGDPKKNVNDTAFTEPMGRPNQDGPALRSIALIKIIDIYIKQSGTDL 227  
 Qy 171 -----YSTATDIWPLVRNDLSVAQYVNTGYDLMEEVNGSSFFTLAVGHRALVEGSA 225  
 Db 228 GAKYKPGSTA-DIFDDIYVRMDLRFLIDHNNSSGDFLMEVNGMHFFLLVQLSAVDSLS 286  
 Qy 226 PATAVSGSCWCD---SQAPRILCYL---QSFWTGSIILANF-----DSSRSGKDA 270  
 Db 287 YFNASERSSPFVEELRQTRDISKFLVDPANGFNGKY--NYIVGPMIADTLRSGLDI 343  
 Qy 271 NTLLGSIHTPDPEAACDSTFQPCSPRALANHKVNDSPFSIYTLNDGLDSEAVAVGRY 330  
 Db 344 STLLAANTVHDAPSA-SHLPFDINDPVLNTLHLLMHMRSIYPIINDSSKNATGIALGRY 402  
 Qy 331 PEDTY-----YNGNPMFLCTIAAEQLYDALYQMDKQSLSEVTVSLDFFKALYSDATG 385  
 Db 403 PEDYVDGYGEGNPMWLATCTASTLYQLYLRHISQHDLVVPMNNDCSNAFMSSELVFS 462  
 Qy 386 TYSS-----SSSTYSSIVDAVKTRPADGVSIIVETHAASNGMSQYDKSDQEQ 433  
 Db 463 NLTLGNDGYLLIEFNTPAFNQTIQKIFQLADSFVLVKAHVGTDGELSQFNRKYYGFW 522  
 Qy 434 LSARDLFWSYAALLTANNRRNSVY 457  
 Db 523 QGAQHLTMSYTSFMDAYQIROEVL 546

## RESULT 15

## ALBYG

glucan 1,4-alpha-glucosidase (BC 3.2.1.3) precursor - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: 1,4-alpha-D-glucan glucohydrolase; glucamylase  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 05-Sep-1997  
 C:Accession: A21896; A23470

R:Yamaashita, I.; Suzuki, K.; Fukui, S.  
 J. Bacteriol. 161, 567-573, 1985

A:Title: Nucleotide sequence of the extracellular glucamylase gene STA1 in the yeast Sa  
 A:Reference number: A21896; MUID:85104778; PMID:3918017  
 A:Note: S. diastaticus

A:Accession: A21896  
 A:Molecule type: DNA  
 A:Residues: 1-778 <YAL>  
 A:Experimental source: strain 5106-9A; ATCC 60709  
 R:Yamaashita, I.; Suzuki, K.; Sakuro, F.  
 Agric. Biol. Chem. 50, 475-482, 1986  
 A:Title: Proteolytic processing of glucamylase in the yeast Saccharomyces cerevisiae.  
 A:Reference number: A23470  
 A:Contents: signal sequence cleavage site  
 A:Accession: A23470  
 A:Molecule type: protein  
 A:Residues: 1-65 <YAZ>  
 A:Gene: STA1  
 C:Superfamily: yeast glucan 1,4-alpha-glucosidase STA1; glucan 1,4-alpha-glucosidase hom  
 C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide c  
 F:1-32/Domain: signal sequence #status experimental <SIG>  
 F:33-778/Product: glucan 1,4-alpha-glucosidase #status predicted <MPT>  
 F:345-778/Domain: glucan 1,4-alpha-glucosidase homology <GAG>  
 F:46,319,333,425,434,445,524,557,656,661,731,752/Binding site: carbohydrate (asn) (coval  
 F:635/Active site: Asp #status predicted

Query Match 18.8%; Score 522; DB 1; Length 778;  
 Best Local Similarity 30.5%; Pred. No. 2.9e-27;  
 Matches 139; Conservative 80; Mismatches 177; Indels 60; Gaps 15;

Qy 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVSGADSGI 60  
 Db 324 VQLADAVLNGTGYVD-SNGAMDSALBWLQCKKYSIRIFENIGPSAVYPS-ILPGV 381  
 Qy 61 VVASPSTNDPDYFTYTRDGLVLTVDLFRNGDTSLSSTIENYSAQAIIVQGISNPSG 120  
 Db 382 VIAPSPQTHPDYFYQWIRDSALTTINSIVS--HSADPA-IEYLLQYLVNSFHLQRTNN--- 435  
 Qy 121 DLSSGAG-----LGEPRNVDEATYSGRPGQPDGAPALRATAMIGFGWLLDNG--- 170  
 Db 436 TLGAGIGYNTDVTALGDPKKNVNDTAFTEPMGRPNQDGPALRSIALIKIIDIYIKQSGTDL 495  
 Qy 171 -----YSTATDIWPLVRNDLSVAQYVNTGYDLMEEVNGSSFFTLAVGHRALVEGSA 225  
 Db 496 GAKYKPGSTA-DIFDDIYVRMDLRFLIDHNNSSGDFLMEVNGMHFFLLVQLSAVDSLS 554  
 Qy 226 PATAVSGSCWCD---SQAPRILCYL---QSFWTGSIILANF-----DSSRSGKDA 270  
 Db 555 YFNASERSSPFVEELRQTRDISKFLVDPANGFNGKY--NYIVGPMIADTLRSGLDI 611  
 Qy 271 NTLLGSIHTPDPEAACDSTFQPCSPRALANHKVNDSPFSIYTLNDGLDSEAVAVGRY 330  
 Db 612 STLLAANTVHDAPSA-SHLPFDINDPVLNTLHLLMHMRSIYPIINDSSKNATGIALGRY 670  
 Qy 331 PEDTY-----YNGNPMFLCTIAAEQLYDALYQMDKQSLSEVTVSLDFFKALYSDATG 385  
 Db 671 PEDYVDGYGEGNPMWLATCTASTLYQLYLRHISQHDLVVPMNNDCSNAFMSSELVFS 730  
 Qy 386 TYSS-----SSSTYSSIVDAVKTRPADGV 409  
 Db 731 NLTLGNDGYLLIEFNTPAFNQTIQKIFQLADSFVL 766

Search completed: June 17, 2004, 17:48:00  
 Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:35:48 ; Search time 59 Seconds  
(without alignment)  
2557.294 Million cell updates/sec

Title: US-10-038-723-2  
Perfect score: 2771  
Sequence: 1 MSPSLALSLGLVCTGLANV.....SKTTATASKTSTTTSGMSL 534

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A.Geneseq.29Jan04:\*

1: geneseqp1980a:\*  
2: geneseqp1990a:\*  
3: geneseqp2000a:\*  
4: geneseqp2001a:\*  
5: geneseqp2002a:\*  
6: geneseqp2003a:\*  
7: geneseqp2003b:\*  
8: geneseqp2004a:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2771	100.0	534	2	AAV18090 Truncated
2	2771	100.0	534	2	AAV23338
3	2771	100.0	534	3	AAV23337
4	2771	100.0	534	3	AAV23337
5	2771	100.0	534	4	AAV7740 A. niger
6	2771	100.0	534	4	AAV7740 A. niger
7	2742	99.0	640	4	AAV7741 A. niger
8	2742	99.0	640	4	AAV7741 A. niger
9	2683.5	96.8	639	1	AAV18090 A. niger
10	2658	95.9	630	1	AAV18090 A. niger
11	2628	94.8	616	3	AAV15176 A. niger
12	2625	94.7	616	3	AAV15176 A. niger
13	2625	94.7	616	3	AAV15176 A. niger
14	2615	94.4	616	2	AAV55976 A. niger
15	2615	94.4	616	2	AAV55976 A. niger
16	2614	94.3	616	3	AAV15184 A. niger
17	2612	94.3	616	2	AAV55980 A. niger
18	2612	94.3	616	2	AAV55980 A. niger
19	2604	94.0	616	2	AAV55977 A. niger
20	2604	94.0	616	2	AAV55977 A. niger
21	2600.5	93.8	621	2	AAV55978 A. niger
22	2589.5	93.5	621	3	AAV15183 A. niger
23	2584.5	93.3	621	3	AAV15183 A. niger
24	2523	89.1	1095	6	ABP96630 A. niger
25	2468	89.1	471	4	ABM00045 AMG SEQ I

26	1850.5	66.8	631	6	ABB80181 A. fumiga
27	1770	63.9	624	4	AAV23339 Talaromyces
28	1589	61.0	618	2	AAV23337 Talaromyces
29	1674.5	60.4	591	2	AAV23337 Talaromyces
30	1615.5	58.3	630	5	AAV51596 Thielavia
31	1577	56.9	626	2	AAV71034 A. niger
32	1425.5	51.4	581	3	AAV18823 Amino acid
33	1425.5	51.4	581	5	AAV79444 Fusarium
34	1425.5	51.4	581	7	ABO01925 A. niger
35	1408.5	50.8	704	6	ABB80170 A. fumiga
36	1380.5	49.8	620	6	ABB80174 A. fumiga
37	1321	47.7	616	6	AAV30155 Glucoamylase
38	725	26.2	579	6	ABP96632 A. niger
39	724	26.1	604	1	AAV60359 Glucoamylase
40	641.5	22.2	624	2	AAV77674 Glucoamylase
41	615.5	22.2	497	1	AAV70572 Glucoamylase
42	600.5	21.7	918	1	AAV60723 Sequence
43	374	13.5	293	1	AAV70183 Sequence
44	181	6.5	1588	7	ADC01413 Enterohae
45	172	6.2	32	2	AAV22343 His(184)

## ALIGNMENTS

## RESULT 1

ID AAV18090 standard; protein; 534 AA.

XX AAV18090;

XX 06-AUG-1999 (first entry)

DE Truncated A. niger glucoamylase GI protein sequence.

XX Glucoamylase; saccharification; liquefied starch solution; cyclodextrin;

XX enzymatic saccharification; monosaccharide production; oligosaccharide;

XX dextrose; trehalose; isomaltoligosaccharide; maltoligosaccharide.

XX Aspergillus niger.

XX W09927124-A1.

XX 03-JUN-1999.

XX 23-NOV-1998; 98MO-US024871.

XX 26-NOV-1997; 97US-00979673.

XX 30-JUN-1998; 98US-00107657.

XX (NOVO ) NOVO-NORDISK AS.

XX (STAL ) STALEY MFG CO A E.

XX Liaw GC, Pedersen S, Hendriksen HV, Svendsen A, Nielsen BR;

XX Nielsen RI.

XX WPI; 1999-357849/30.

XX N-PSDB; AAV76983.

XX Saccharification of liquefied starch solution for production of, e.g.

XX monosaccharides.

XX Claim 18; Page 55-58; 63pp; English.

XX This sequence is a truncated A. niger glucoamylase GI protein. The

XX invention relates to a method of saccharifying a liquefied starch

XX solution comprising: (a) a saccharification step during which one or more

XX enzymatic saccharification stages (using a mutated Aspergillus niger

XX glucoamylase) take place; (b) one or more high temperature membrane

XX separation steps; and (c) re-circulation of the saccharification enzyme;

XX in which the membrane separation steps are carried out as an integral

XX part of the saccharification step. The method can be used for the

XX production of mono and/or oligosaccharides from starch, including

CC dextrose, trehalose, isomaltotriosesaccharides, cyclodextrins or  
CC maltotriosesaccharides. The use of the membrane separation step improves  
CC the efficiency of the process and improves yields and purity  
XX

XX Sequence 534 AA:

Query Match 100.0%; Score 2771; DB 2; Length 534;  
Best Local Similarity 100.0%; Pred. No. 1.5e-214;  
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSFSLIALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVSGADSGI 60
DB 1 MSFSLIALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVSGADSGI 60
QY 61 VVASPTDNDPYFTYTWTRDSGLVTKTVLDLFRNGDTSLSSTIENYISAQAIIVOGISNPSG 120
DB 61 VVASPTDNDPYFTYTWTRDSGLVTKTVLDLFRNGDTSLSSTIENYISAQAIIVOGISNPSG 120
QY 121 DLSGAGLGEPRKFNVDETATYTGSGWRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
DB 121 DLSGAGLGEPRKFNVDETATYTGSGWRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFETIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
DB 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFETIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
QY 241 APEILCYLOSFWTGSFILANFDSRSRGKANTLLGSIHTFDPEACDSTFQPCSPRALA 300
DB 241 APEILCYLOSFWTGSFILANFDSRSRGKANTLLGSIHTFDPEACDSTFQPCSPRALA 300
QY 301 NHKEVDSFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLAABQLYDALYQMD 360
DB 301 NHKEVDSFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLAABQLYDALYQMD 360
QY 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAKTPADGVSIIVETHAASNG 420
DB 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAKTPADGVSIIVETHAASNG 420
QY 421 SMSQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVGTCAATSAIG 480
DB 421 SMSQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVGTCAATSAIG 480
QY 481 TYSSTVTSWPSIVATGTTTATPTGSGSVTSKTTATASKTSTTRGMSL 534
DB 481 TYSSTVTSWPSIVATGTTTATPTGSGSVTSKTTATASKTSTTRGMSL 534
```

RESULT 2

AA2338 ID AA2338 standard; protein; 534 AA.

XX AA2338;

XX 02-SEP-1999 (first entry)

XX Aspergillus niger glucoamylase enzyme.

XX Glucoamylase enzyme; saccharification; starch hydrolyase;

XX dextrose syrup; syrup; ethanol; fuel; beverage; citric acid;

XX ascorbic acid; lysine; glutamic acid.

XX Aspergillus niger.

XX MO928448-A1.

XX 10-JUN-1999.

XX 26-NOV-1998; 98W0-DK000520.

XX 26-NOV-1997; 97US-00979673.

XX 30-DEC-1997; 97DK-00001557.

XX 30-JUN-1998; 98US-00107657.

XX 10-JUN-1998; 98DK-00000925.

XX (NOVO ) NOVO-NORDISK AS.

XX Nielsen BR, Nielsen RI, Lehmbeck J;

XX WPI; 1999-404822/34.

XX N-PSDB; AAX81797.

XX New glucoamylase obtained from Talaromyces emersonii.

XX Disclosure; Page 63-64; 79pp; English.

XX The present sequence represents a glucoamylase enzyme obtained from  
XX Aspergillus niger. The specification describes a Talaromyces emersonii  
XX glucoamylase enzyme. The glucoamylase enzymes have high thermal stability  
XX so that a saccharification process may be carried out within a shorter  
XX period of time or the process may be carried out using a lower enzyme  
XX dosage. The glucoamylase enzymes can be used for saccharifying starch  
XX hydrolyzate for converting starch or partially hydrolysed starch into a  
XX syrup containing dextrose. They can be used for producing  
XX oligosaccharides, specialty syrups, ethanol for fuel, beverages or  
XX organic compounds such as citric acid, ascorbic acid, lysine or glutamic  
XX acid

XX Sequence 534 AA;

Query Match 100.0%; Score 2771; DB 2; Length 534;  
Best Local Similarity 100.0%; Pred. No. 1.5e-214;  
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSFSLIALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVSGADSGI 60
DB 1 MSFSLIALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVSGADSGI 60
QY 61 VVASPTDNDPYFTYTWTRDSGLVTKTVLDLFRNGDTSLSSTIENYISAQAIIVOGISNPSG 120
DB 61 VVASPTDNDPYFTYTWTRDSGLVTKTVLDLFRNGDTSLSSTIENYISAQAIIVOGISNPSG 120
QY 121 DLSGAGLGEPRKFNVDETATYTGSGWRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
DB 121 DLSGAGLGEPRKFNVDETATYTGSGWRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFETIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
DB 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFETIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
QY 241 APEILCYLOSFWTGSFILANFDSRSRGKANTLLGSIHTFDPEACDSTFQPCSPRALA 300
DB 241 APEILCYLOSFWTGSFILANFDSRSRGKANTLLGSIHTFDPEACDSTFQPCSPRALA 300
QY 301 NHKEVDSFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLAABQLYDALYQMD 360
DB 301 NHKEVDSFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLAABQLYDALYQMD 360
QY 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAKTPADGVSIIVETHAASNG 420
DB 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAKTPADGVSIIVETHAASNG 420
QY 421 SMSQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVGTCAATSAIG 480
DB 421 SMSQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVGTCAATSAIG 480
QY 481 TYSSTVTSWPSIVATGTTTATPTGSGSVTSKTTATASKTSTTRGMSL 534
DB 481 TYSSTVTSWPSIVATGTTTATPTGSGSVTSKTTATASKTSTTRGMSL 534
```

RESULT 3

AAB03450 ID AAB03450 standard; protein; 534 AA.

XX AAB03450;

DT 03-JAN-2001 (first entry)  
 XX Aspergillus niger G2 glucosylase.  
 XX Glucosylase G2; starch hydrolysis; high fructose corn syrup;  
 KM thermal stability.  
 XX Aspergillus niger.  
 XX Key Location/Qualifiers  
 FT Peptide 1..27  
 FT /label= prepropeptide\_sequence  
 FT Cleavage-site 27..28  
 FT /label= tripeptidyl\_aminopeptidase\_cleavage\_site  
 FT Protein 26..534  
 FT /label= mature\_glucosylase  
 XX WO200034452-A1.  
 XX 15-JUN-2000.  
 XX 07-DEC-1999; 99WO-DK000666.  
 XX 07-DEC-1998; 98DK-00001616.  
 PR 24-MAR-1999; 99DK-00000409.  
 XX (NOVO ) NOVO-NORDISK AS.  
 XX Nielsen BR, Svendsen A, Boisen K, Vind J, Pedersen H;  
 PI MPI; 2000-431296/37.  
 DR Variants of parent fungal glucosylase useful for producing ethanol,  
 PT beverages and in fermentation processes comprise a peptide extension at  
 PT the amino terminal.  
 XX Disclosure; Page 50-51; 61pp; English.  
 XX The present sequence is the G2 glucosylase from Aspergillus niger. It  
 CC was used to construct a version of the enzyme which has improved thermal  
 CC stability, and which can be used more efficiently in starch hydrolysis.  
 CC This process is used during high fructose corn syrup production, as well  
 CC as in the production of ethanol for fuel or beverage, and in fermentation  
 CC processes for producing organic compounds such as citric acid, ascorbic  
 CC acid, lysine and glutamic acid  
 CC  
 XX Sequence 534 AA;  
 SQ  
 Query Match 100.0%; Score 2771; DB 3; Length 534;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;  
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 301 NHKEVVDSPFRSITYLTNDGLSDSEAVAVGRYPEDTYNNGNPFLLCTLAAREQLYDALYQWD 360  
 QY 361 KQGSLEVTDVSLDPFKALYSDAATGTYSSSSSTYSIDAVKTPDGPVSIYETAASNG 420  
 DB 361 KQGSLEVTDVSLDPFKALYSDAATGTYSSSSSTYSIDAVKTPDGPVSIYETAASNG 420  
 QY 421 SMSSEYDSDGSEQLSARDLUTWSYALLTANNRRNSVVPASMGETSASSVPGCAATSAIG 480  
 DB 421 SMSSEYDSDGSEQLSARDLUTWSYALLTANNRRNSVVPASMGETSASSVPGCAATSAIG 480  
 QY 481 TYSSVTVTSWPSIYATGTTTATPTPGSGSVTSISKTATASKTSTTRSGMSL 534  
 DB 481 TYSSVTVTSWPSIYATGTTTATPTPGSGSVTSISKTATASKTSTTRSGMSL 534  
 RESULT 4  
 AA77740  
 ID AA77740 standard; protein; 534 AA.  
 XX  
 AC AA77740;  
 XX  
 DT 22-MAY-2000 (first entry)  
 XX  
 DE A. niger G2 glucosylase.  
 XX  
 KM Glucosylase; variant; starch conversion; saccharification; ethanol;  
 KM fuel; beverage; fermentation; citric acid; ascorbic acid; thermostable;  
 KM G2 glucosylase; fungal.  
 XX  
 OS Aspergillus niger.  
 XX  
 PN WO200004136-A1.  
 PD 27-JAN-2000.  
 XX  
 PF 09-JUL-1999; 99WO-DK000392.  
 XX  
 PR 15-JUL-1998; 98DK-00000937.  
 PR 17-DEC-1998; 98DK-00001667.  
 XX  
 PA (NOVO ) NOVO-NORDISK AS.  
 XX  
 PI Nielsen BR, Svendsen A, Pedersen H, Vind J, Hendriksen HV;  
 PI Prandsen TP;  
 DR MPI; 2000-182412/16.  
 DR N-PSDB; AA287842.  
 XX  
 PT Variant fungal glucosylases with improved thermostability and increased  
 PT specific activity, useful in saccharification processes.  
 XX  
 PS Claim 1; Page 78-81; 116pp; English.  
 XX  
 CC The invention relates to variant fungal glucosylases comprising specific  
 CC mutations. The variant comprises one or more mutation in position/region  
 CC 1-18, 19-35, 40-62, 73-80, 93-127, 170-184, 200-212, 234-246, 287-319,  
 CC 334-341, 353-374, 388-414, 445-470 of the parent G2 glucosylase sequence  
 CC (AA77740) from A. niger, with the exception of: N20C, A27C, S30P, Y48W,  
 CC W50H, W52P, R54K/L, D55G/V, G57A, K108R, D112Y, Y116A/W, S119C/W/E/G/Y/P,  
 CC W120H/L/F/Y, G121T/A, R122Y, P123G, Q124H, R125K, W170F, N175S, Q172N, CC  
 CC T172G, G174C, Y175F, D176N/E, L177H/D, W178R/D, E179Q/D, E180D/O,  
 CC V181D/A/T, N182A/D/Q/Y/S, G183K, S184H, W212F, R241K, A246C, D253E/Q,  
 CC A303V, R305K, Y306F, D309N/E, Y312W, W317F, E389D/O, H391W, A392D, A393P,  
 CC N395Q, G396S, E400Q/C, Q401E, G407D, E408P, L410F, S411A/G/C/H/D, and  
 CC S460P. The glucosylase variants are useful in a starch conversion  
 CC process, especially continuous processes which include a continuous  
 CC saccharification process. The variants can be used for producing  
 CC oligosaccharides, especially syrups, or ethanol for fuel or beverages.  
 CC They can also be used in fermentation processes for producing organic  
 CC compounds such as citric acid, ascorbic acid, lysine and glutamic acid.  
 CC The glucosylase variants have improved thermostability and/or increased  
 CC specific activity. This is advantageous in industrial saccharification

CC processes. The risk of microbial contamination is also reduced when  
 CC carrying the saccharification process at temperatures above 63 plusOC. An  
 CC increased specific activity towards short chain saccharides such as  
 CC maltose (without reducing the activity towards oligosaccharides) would  
 CC also permit using a lower enzyme dosage and/or shorter process times. The  
 CC present sequence represents the G2 glucanase from *A. niger*, the parent  
 CC glucanase used for constructing the variants

XX Sequence 534 AA;

Query Match 100.0%; Score 2771; DB 3; Length 534;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;  
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFRLALSGVCTGLANVSKRATLDSWLSNEATVARTAIINNIGADGAWGADSGI 60  
 DB 1 MSFRLALSGVCTGLANVSKRATLDSWLSNEATVARTAIINNIGADGAWGADSGI 60  
 QY 61 VVASPSTDNPDYFYTWTRDSGLVLTVDLFRNGDTSLSSTIENYISAQAIIVOGISNPSG 120  
 DB 61 VVASPSTDNPDYFYTWTRDSGLVLTVDLFRNGDTSLSSTIENYISAQAIIVOGISNPSG 120  
 QY 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALRATAMIGFGQWLLDNGYSTATDIW 180  
 DB 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALRATAMIGFGQWLLDNGYSTATDIW 180  
 QY 181 PLYVNDLSYVAQYWNQGYDLMEVNGSSFFTIAYQHRALVEGSAFATAVSSCSWCDQ 240  
 DB 181 PLYVNDLSYVAQYWNQGYDLMEVNGSSFFTIAYQHRALVEGSAFATAVSSCSWCDQ 240  
 QY 241 APEILCYLQSFMTGSFILLNFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPALA 300  
 DB 241 APEILCYLQSFMTGSFILLNFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPALA 300  
 QY 301 NHKEVDSFESIYTLNDGLSDSEAVAGRPEDTYNGNPMFLCTLAABQLYDALYQWD 360  
 DB 301 NHKEVDSFESIYTLNDGLSDSEAVAGRPEDTYNGNPMFLCTLAABQLYDALYQWD 360  
 QY 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAKTPADGFSIVETHAASNG 420  
 DB 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAKTPADGFSIVETHAASNG 420  
 QY 421 SMSFOYKSDGEQLSARDLTWSYAAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480  
 DB 421 SMSFOYKSDGEQLSARDLTWSYAAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480  
 QY 481 TYSSVTVTWSPIVATGTTTATPTGSGVSTSTSKTTATASKTSTTRSGMSL 534  
 DB 481 TYSSVTVTWSPIVATGTTTATPTGSGVSTSTSKTTATASKTSTTRSGMSL 534

RESULT 5  
 AAB48171 standard; procein; 534 AA.

XX AAB48171;

XX 02-APR-2001 (first entry)

XX A. niger G1 glucanase polypeptide.

XX Thermoascus crustaceus; glucanase; starch conversion; ethanol;  
 KW maltose syrup; beverage; citric acid; ascorbic acid; detergent;  
 KM thermostability; glucose; G1 glucanase.

XX Aspergillus niger.

XX Key Location/Qualifiers

XX Peptide 1..24

XX Protein /note= "signal peptide"  
 25..534

XX /note= "mature protein"

PN W0200075296-A1.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000MO-DK000301.

XX 02-JUN-1999; 99DK-00000779.

XX (NOVO ) NOVO NORDISK AS.

PI Nielsen BR, Kauppinen MS, Nielsen RI;

XX WPI; 2001-071066/08.

DR N-PSDB; AAC84444.

PT Novel Thermoascus crustaceus glucanase polypeptide useful in a  
 PT continuous starch conversion process, detergents and for producing  
 PT oligosaccharides, specially syrups, ethanol for fuel or drinking, and  
 PT beverages.

PS Disclosure; Page 52-55; 61pp; English.

CC The invention provides a Thermoascus crustaceus glucanase polypeptide.  
 CC The glucanase polypeptide can be expressed by standard recombinant  
 CC methodology and is useful in a continuous starch conversion process, for  
 CC producing oligosaccharides, specially syrups such as maltose syrups,  
 CC ethanol for fuel or drinking ethanol, beverages, and organic compounds  
 CC such as citric acid, ascorbic acid, lysine or glutamic acid. It is also  
 CC useful in detergents such as laundry detergent compositions, dish wash  
 CC compositions and/or hard surface cleaning compositions. The T. crustaceus  
 CC glucanase has higher thermostability than *Aspergillus niger* G1  
 CC glucanase. It also has higher specific activity and/or decreased  
 CC glucose reversion tendency. The present sequence represents the *A. niger*  
 CC G1 glucanase

XX Sequence 534 AA;

Query Match 100.0%; Score 2771; DB 4; Length 534;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;  
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFRLALSGVCTGLANVSKRATLDSWLSNEATVARTAIINNIGADGAWGADSGI 60  
 DB 1 MSFRLALSGVCTGLANVSKRATLDSWLSNEATVARTAIINNIGADGAWGADSGI 60  
 QY 61 VVASPSTDNPDYFYTWTRDSGLVLTVDLFRNGDTSLSSTIENYISAQAIIVOGISNPSG 120  
 DB 61 VVASPSTDNPDYFYTWTRDSGLVLTVDLFRNGDTSLSSTIENYISAQAIIVOGISNPSG 120  
 QY 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALRATAMIGFGQWLLDNGYSTATDIW 180  
 DB 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALRATAMIGFGQWLLDNGYSTATDIW 180  
 QY 181 PLYVNDLSYVAQYWNQGYDLMEVNGSSFFTIAYQHRALVEGSAFATAVSSCSWCDQ 240  
 DB 181 PLYVNDLSYVAQYWNQGYDLMEVNGSSFFTIAYQHRALVEGSAFATAVSSCSWCDQ 240  
 QY 241 APEILCYLQSFMTGSFILLNFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPALA 300  
 DB 241 APEILCYLQSFMTGSFILLNFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPALA 300  
 QY 301 NHKEVDSFESIYTLNDGLSDSEAVAGRPEDTYNGNPMFLCTLAABQLYDALYQWD 360  
 DB 301 NHKEVDSFESIYTLNDGLSDSEAVAGRPEDTYNGNPMFLCTLAABQLYDALYQWD 360  
 QY 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAKTPADGFSIVETHAASNG 420  
 DB 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAKTPADGFSIVETHAASNG 420  
 QY 421 SMSFOYKSDGEQLSARDLTWSYAAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480  
 DB 421 SMSFOYKSDGEQLSARDLTWSYAAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480

QY 481 TYSSTVTSMPISIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534  
 DB 481 TYSSTVTSMPISIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534

RESULT 6  
 AAB61904  
 ID AAB61904 standard; protein; 534 AA.  
 AC AAB61904;  
 DT 08-MAY-2001 (first entry)  
 DE A. niger G2 glucoamylase.

KM Glucoamylase; variant; starch; dextrose; saccharification; maltodextrin;  
 KW glucose syrup; fuel; ethanol; beverage; fermentation; citric acid;  
 KM ascorbic acid.  
 OS Aspergillus niger.

FT Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /note= "signal peptide"  
 FT Protein 25..534  
 FT /note= "mature protein"

MO200104273-A2.  
 PD 18-JAN-2001.  
 PF 07-JUL-2000; 2000MO-DK000373.  
 PR 09-JUL-1999; 99DK-00000999.

XX (NOVO ) NOVO NORDISK AS.

XX PI Nielsen BR, Svendsen A, Pedersen H, Vind J, Hendriksen HV,  
 XX Frandsen TP;

XX DR WPI; 2001-138334/14.  
 XX N-PSDB; AAC85099.

PT Novel variant of parent glucoamylase useful in starch conversion process,  
 PT and for producing oligosaccharides, maltodextrins, glucose syrups, fuel,  
 PT drinking ethanol, beverage and organic compounds.

PS Claim 1; Page 51-53; 58pp; English.

XX The invention relates to a variant of a parent glucoamylase, comprising  
 CC an alteration at positions 59, 66, 72, 119, 189, 223, 227, 313, 340, 342,  
 CC 352, 379, 386, 393, 395, 402, 408, 416, 425, 427, 444, 486, 490, 494,  
 CC where the alteration is independently an insertion, substitution or  
 CC deletion of amino acid which occupies the position. The variant  
 CC glucoamylase is useful for converting starch or partially hydrolyzed  
 CC starch into a syrup containing dextrose, by saccharifying starch  
 CC hydrolyzate. The variant is useful in the starch conversion process, for  
 CC producing oligosaccharides, maltodextrins or glucose syrups, fuel,  
 CC drinking ethanol, beverage and in a fermentation process for producing  
 CC organic compounds, such as citric acid, ascorbic acid, lysine and  
 CC glutamic acid. It is useful for improving the thermal stability and/or  
 CC specific activity of a parent glucoamylase. The present sequence  
 CC represents an A. niger G2 glucoamylase, the parent enzyme from which the  
 CC variant of the invention is derived

XX Sequence 534 AA;

Query Match 100.0%; Score 2771; DB 4; Length 534;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;  
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFRLSLALSGVCTGLANVISKRATLDLWLSNEATVARTAILNNIGADGAWGADSGI 60  
 PS ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1 MSFRLSLALSGVCTGLANVISKRATLDLWLSNEATVARTAILNNIGADGAWGADSGI 60  
 QY 61 VVASSTSTNDPQFYMTWTDGSLVLTLYDLFENGSTSLSTENTYISAQAIYQGISNPSG 120  
 DB 61 VVASSTSTNDPQFYMTWTDGSLVLTLYDLFENGSTSLSTENTYISAQAIYQGISNPSG 120  
 QY 121 DLSGAGLGEPEFNVDETAFTYSGMGRPORDGALBATAIMIGEQMLDNGYSTATDIYW 180  
 DB 121 DLSGAGLGEPEFNVDETAFTYSGMGRPORDGALBATAIMIGEQMLDNGYSTATDIYW 180  
 QY 181 PLVRNDLSYVAQYMNQGTGYDLMEBVNGSSFPTIAVQHRALYEGSAFATVAGSSCSWCDSQ 240  
 DB 181 PLVRNDLSYVAQYMNQGTGYDLMEBVNGSSFPTIAVQHRALYEGSAFATVAGSSCSWCDSQ 240  
 QY 241 APELICYSQFWTGSFILANPSSSGDANTLLGSIHTPDEACDOSTPQCPSPRALA 300  
 DB 241 APELICYSQFWTGSFILANPSSSGDANTLLGSIHTPDEACDOSTPQCPSPRALA 300  
 QY 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTILAAAEQLYDALYQMD 360  
 DB 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTILAAAEQLYDALYQMD 360  
 QY 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFPVSIYETHAASNG 420  
 DB 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFPVSIYETHAASNG 420  
 QY 421 SMSQYDYSDEQSLASBDLTWSYALLTLNRRNRNVVPASMETASSVPGCAATSAIG 480  
 DB 421 SMSQYDYSDEQSLASBDLTWSYALLTLNRRNRNVVPASMETASSVPGCAATSAIG 480  
 QY 481 TYSSTVTSMPISIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534  
 DB 481 TYSSTVTSMPISIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534

RESULT 7  
 AAY77741

ID AAY77741 standard; protein; 640 AA.

XX AAY77741;

XX DT 22-MAY-2000 (first entry)

XX DE A. niger G1 glucoamylase.

XX KM Glucoamylase; variant; starch conversion; saccharification; ethanol;  
 KW fuel; beverage; fermentation; citric acid; ascorbic acid; thermostable;  
 XX G1 glucoamylase; fungal.

XX OS Aspergillus niger.

XX PN WO200004136-A1.

XX PD 27-JAN-2000.

XX PF 09-JUL-1999; 99MO-DK000392.

XX PR 15-JUL-1998; 98DK-00000937.

XX PR 17-DEC-1998; 98DK-00001667.

XX PA (NOVO ) NOVO-NORDISK AS.

XX PI Nielsen BR, Svendsen A, Pedersen H, Vind J, Hendriksen HV,  
 XX Frandsen TP;

XX DR WPI; 2000-182412/16.  
 XX N-PSDB; AA87843.

PT Variant fungal glucoamylases with improved thermostability and increased  
 PT specific activity, useful in saccharification processes.

PS Disclosure; Page 91-93; 116pp; English.

XX



CC The invention relates to variant fungal glucosylases. The variants  
CC comprise specific mutations in the parent G2 glucosylase (Amg) sequence  
CC (AAV77740) from *A. niger* (see AA287842 for specific positions of the  
CC mutations). The glucosylase variants are useful in a starch conversion  
CC process, especially continuous process which include a continuous  
CC saccharification process. The variants can be used for producing  
CC oligosaccharides, especially syrups, or ethanol for fuel or beverages.  
CC They can also be used in fermentation processes for producing organic  
CC compounds such as citric acid, ascorbic acid, lysine and glutamic acid.  
CC The glucosylase variants have improved thermostability and/or increased  
CC specific activity. This is advantageous in industrial saccharification  
CC processes. The risk of microbial contamination is also reduced when  
CC carrying the saccharification process at temperatures above 63 plusOC. An  
CC increased specific activity towards short chain saccharides such as  
CC maltose (without reducing the activity towards oligosaccharides) would  
CC also permit using a lower enzyme dosage and/or shorter process times. The  
CC present sequence represents the G1 glucosylase from *A. niger*  
XX  
SQ Sequence 640 AA;  
Query Match 99.0%; Score 2742; DB 3; Length 640;  
Best Local Similarity 99.6%; Pred. No. 4.4e-212;  
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWSGADSGI 60  
DB 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWSGADSGI 60  
QY 61 VVASPTDNDPXYFTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120  
DB 61 VVASPTDNDPXYFTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120  
QY 121 DLSSGAGLGEKPFVNDATAVYGSWGRPQRDPALRATAMIFGQWLLDNGTSTATDIW 180  
DB 121 DLSSGAGLGEKPFVNDATAVYGSWGRPQRDPALRATAMIFGQWLLDNGTSTATDIW 180  
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNCSFFTIAVQHRALVEGSAFATAVSSCSWCDQ 240  
DB 181 PLVNDLSYVAQYNNQGYDLMEEVNCSFFTIAVQHRALVEGSAFATAVSSCSWCDQ 240  
QY 241 APEILCYLQSFMTGSFILLANFDSRSRSGKANTLLGSIHTFPEAACDSTFQPCSPRALA 300  
DB 241 APEILCYLQSFMTGSFILLANFDSRSRSGKANTLLGSIHTFPEAACDSTFQPCSPRALA 300  
QY 301 NHKEVDSFRSITLTLNDGLDSEAVAVGRYPEDTYNNGNPMFLCTLAABQLYDALYQMD 360  
DB 301 NHKEVDSFRSITLTLNDGLDSEAVAVGRYPEDTYNNGNPMFLCTLAABQLYDALYQMD 360  
QY 361 KQSLLEVTVDSLDFFKALYSDAATGYSSSSSTYSIVDAVKTFADGFSIVETHAASNG 420  
DB 361 KQSLLEVTVDSLDFFKALYSDAATGYSSSSSTYSIVDAVKTFADGFSIVETHAASNG 420  
QY 421 SMSOYDSDGEOLASRDLTWSYALLTANNRNSVVPASGCTSSASVPTCAATSAIG 480  
DB 421 SMSOYDSDGEOLASRDLTWSYALLTANNRNSVVPASGCTSSASVPTCAATSAIG 480  
QY 481 TYSSVTVSWPSIVATGCTTTATPTGSGVTSKTTATASKTSTTSS 530  
DB 481 TYSSVTVSWPSIVATGCTTTATPTGSGVTSKTTATASKTSTTSS 530  
RESULT 8  
AAB61905  
ID AAB61905 standard; protein; 640 AA.  
XX  
AC AAB61905;  
XX  
DT 08-MAY-2001 (first entry)  
DE  
XX A. niger protein sequence Id No. 13.  
XX Glucoamylase; variant; starch; dextrose; saccharification; maltodextrin;  
KW glucose syrup; fuel; ethanol; beverage; fermentation; citric acid;

KW ascorbic acid.  
XX  
XX Aspergillus niger.  
OS  
XX WO200104273-A2.  
PN  
XX 18-JAN-2001.  
PD  
XX 07-JUL-2000; 2000WO-DK00373.  
PF  
XX 09-JUL-1999; 99DK-00000999.  
PR  
XX (NOVO ) NOVO NOBDISK AS.  
PA  
XX Nielsen BR, Svendsen A, Pedersen H, Vind J, Hendriksen HV;  
PI Frandsen TP;  
PI  
XX WPI; 2001-138334/14.  
DR  
XX  
XX Novel variant of parent glucosylase useful in starch conversion process,  
PT and for producing oligosaccharides, maltodextrins, glucose syrups, fuel,  
PT drinking ethanol, beverage and organic compounds.  
PS Disclosure; Page 57-58; 58pp; English.  
PS  
XX The invention relates to a variant of a parent glucosylase, comprising  
CC an alteration at positions 59, 66, 72, 119, 189, 223, 227, 313, 340, 342,  
CC 352, 379, 386, 393, 395, 402, 408, 416, 425, 427, 444, 486, 490, 494,  
CC where the alteration is independently an insertion, substitution or  
CC deletion of amino acid which occupies the position. The variant  
CC glucosylase is useful for converting starch or partially hydrolyzed  
CC starch into a syrup containing dextrose, by saccharifying starch  
CC hydrolyzate. The variant is useful in the starch conversion process, for  
CC producing oligosaccharides, maltodextrins or glucose syrups, fuel,  
CC drinking ethanol, beverage and in a fermentation process for producing  
CC organic compounds, such as citric acid, ascorbic acid, lysine and  
CC glutamic acid. It is useful for improving the thermal stability and/or  
CC specific activity of a parent glucosylase  
XX  
SQ Sequence 640 AA;  
Query Match 99.0%; Score 2742; DB 4; Length 640;  
Best Local Similarity 99.6%; Pred. No. 4.4e-212;  
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWSGADSGI 60  
DB 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWSGADSGI 60  
QY 61 VVASPTDNDPXYFTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120  
DB 61 VVASPTDNDPXYFTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120  
QY 121 DLSSGAGLGEKPFVNDATAVYGSWGRPQRDPALRATAMIFGQWLLDNGTSTATDIW 180  
DB 121 DLSSGAGLGEKPFVNDATAVYGSWGRPQRDPALRATAMIFGQWLLDNGTSTATDIW 180  
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNCSFFTIAVQHRALVEGSAFATAVSSCSWCDQ 240  
DB 181 PLVNDLSYVAQYNNQGYDLMEEVNCSFFTIAVQHRALVEGSAFATAVSSCSWCDQ 240  
QY 241 APEILCYLQSFMTGSFILLANFDSRSRSGKANTLLGSIHTFPEAACDSTFQPCSPRALA 300  
DB 241 APEILCYLQSFMTGSFILLANFDSRSRSGKANTLLGSIHTFPEAACDSTFQPCSPRALA 300  
QY 301 NHKEVDSFRSITLTLNDGLDSEAVAVGRYPEDTYNNGNPMFLCTLAABQLYDALYQMD 360  
DB 301 NHKEVDSFRSITLTLNDGLDSEAVAVGRYPEDTYNNGNPMFLCTLAABQLYDALYQMD 360  
QY 361 KQSLLEVTVDSLDFFKALYSDAATGYSSSSSTYSIVDAVKTFADGFSIVETHAASNG 420  
DB 361 KQSLLEVTVDSLDFFKALYSDAATGYSSSSSTYSIVDAVKTFADGFSIVETHAASNG 420



QY 421 SMSEYDKSDGEQLSARDLTWVSYALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480  
 DB 421 SMSEYDKSDGEQLSARDLTWVSYALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480  
 QY 481 TYSSVTWTSWPSIVATGCTTTATPTGSGSVTSTSKTTATASKSTTTTRS 530  
 DB 481 TYSSVTWTSWPSIVATGCTTTATPTGSGSVTSTSKTTATASKSTTTTRS 530

## RESULT 9

AAp40212  
 ID AAP40212 standard; protein; 639 AA.

XX AAP40212;

AC 25-MAR-2003 (revised)

DT 09-JAN-1992 (first entry)

XX Sequence encoded by A.awamori glucoamylase genomic region.

XX Starch hydrolysis; glucose.

XX Aspergillus awamori.

XX WO8402921-A.

XX 02-AUG-1984.

XX 26-JAN-1984; 84WO-US000122.

XX 28-JAN-1983; 83US-00461920.

XX 20-DEC-1983; 83US-00563941.

XX 27-APR-1987; 87US-00047552.

XX (CETU ) CETUS CORP.

XX (CETU ) CETUS CORP.

XX Nunberg JH, Flatsgaard JE, Innis MA, Gelfand DH, Meade JH;

XX WPI; 1984-201413/32.

XX DR N-PSDB; AAN40166.

XX DNA sequence coding for fungal glucoamylase protein - for expression in

XX yeast etc. for prodn. of the enzyme.

XX Example; Table 1, Page 21-25; 66pp; English.

XX The inventors claim a modified DNA sequence coding for fungal

XX glucoamylase protein or its single or multiple base substitutions

XX deletions, insertions or inversions is new (see AAN40165). It is derived

XX from natural, synthetic or semisynthetic sources and is capable, when

XX correctly combined with a cleaved expression vector, of expressing a non-

XX native protein having glucoamylase activity on transformation of a host

XX organism by the vector. (Updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 639 AA;

QY Query Match 96.8%; Score 2683.5; DB 1; Length 639;

DB Best Local Similarity 98.1%; Pred. No. 2.3e-207;

DB Matches 520; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 1 MSFRLLSLGLVCTGLANVISKRAITLDSWLSNEATVARTAILNIGADGAVSGADSGI 60

DB 1 MSFRLLSLGLVCTGLANVISKRAITLDSWLSNEATVARTAILNIGADGAVSGADSGI 60

QY 61 VVASSTONPDYFYWTWTRDSGLVLTVDLFRNGGTSLSLSTENISQAQIVQGISNPSG 120

DB 61 VVASSTONPDYFYWTWTRDSGLVLTVDLFRNGGTSLSLSTENISQAQIVQGISNPSG 120

QY 121 DLSSAGAGIEPKFNVDETAITGSGMGRPORDDGALPATAIMIGFQWLNDNGYSTATDIW 180

DB 121 DLSSAGAGIEPKFNVDETAITGSGMGRPORDDGALPATAIMIGFQWLNDNGYSTATDIW 180

QY 181 PLVRNDLSVVAQYMQTGDLMEEYNGSSFTTIAVOHRAIVEGSAFATAVGSSCSWCDSD 240  
 DB 180 PLVRNDLSVVAQYMQTGDLMEEYNGSSFTTIAVOHRAIVEGSAFATAVGSSCSWCDSD 239  
 QY 241 APEILCYLQSFMTGSFILLANPDSRSRGDANTLGSHTFDEPAACDDSTFPCSPRALA 300  
 DB 240 APEILCYLQSFMTGSFILLANPDSRSRGDANTLGSHTFDEPAACDDSTFPCSPRALA 299

QY 301 NKEEVDSPFRSITLTNDGLSDSEAVAVGRYPEDDTYNGNPMFLCTLAABEQLYDALYQWD 360

DB 300 NKEEVDSPFRSITLTNDGLSDSEAVAVGRYPEDDTYNGNPMFLCTLAABEQLYDALYQWD 359

QY 361 KOGSLEVTVDLSDFPKALYSDAATGTSYSSSTYSIVDAVTFPDGFSIYETAAASNG 420

DB 360 KOGSLEVTVDLSDFPKALYSDAATGTSYSSSTYSIVDAVTFPDGFSIYETAAASNG 419

QY 421 SMSEYDKSDGEQLSARDLTWVSYALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480

DB 420 SMSEYDKSDGEQLSARDLTWVSYALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 479

QY 481 TYSSVTWTSWPSIVATGCTTTATPTGSGSVTSTSKTTATASKSTTTTRS 530

DB 480 TYSSVTWTSWPSIVATGCTTTATPTGSGSVTSTSKTTATASKSTTTTRS 529

## RESULT 10

AAp81876  
 ID AAP81876 standard; protein; 630 AA.

XX AAP81876;

XX 24-OCT-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 15-MAR-1992 (first entry)

XX Sequence of amylglucosidase.

XX Enzyme; brewing; bread-making; dextrin.

XX Aspergillus niger; NCI 22343.

XX EP260160-A.

XX 16-MAR-1986.

XX 10-JUN-1987; 87EP-00401300.

XX 10-JUN-1986; 86FR-00008387.

XX 13-APR-1987; 87FR-00005207.

XX 13-APR-1987; 87FR-00005208.

XX (TRGE ) TRANSGENE SA.

XX Labat N, Loison G, Lemoine Y;

XX WPI; 1988-072593/11.

XX DR N-PSDB; AAN82019.

XX New DNA block for expressing amyloglucosidase in yeast - contg. gene plus

XX transcription and export sequences, and transformed cells useful in

XX brewing and bread making.

XX Example; Fig 2; 38pp; French.

XX EcoRI-sal I fragments of Aspergillus niger NCI 22343 were cloned in

XX pBR322 and two clones contg. portions of the gene isolated using a pool

XX of three probes-TG282, TG283, TG284- (AAN82014-6) designed on the basis

XX of already published SGB. The clones were designated pTG1830 and 1831. A

XX cDNA bank of A.niger NCI 22343 was screened for amyloglucosidase using

XX probe TG433 (AAN82017) which corresp. to AAs 163-172 of the mature

XX protein (see AAN82019). Oligo TG387 (AAN82018) was used to confirm the

XX identity of the clones detected. (Updated on 25-MAR-2003 to correct PR

CC field.) (updated on 25-MAR-2003 to correct PA field.) (updated on 24-OCT-  
CC 2003 to standardise OS field)  
XX  
SQ Sequence 630 AA;  
Query Match 95.9%; Score 2658; DB 1; Length 630;  
Best Local Similarity 97.2%; Pred. No. 2,6e-205;  
Matches 515; Conservative 2; Mismatches 3; Indels 10; Gaps 1;  
QY 1 MSFSLALSGVCTGLANVSKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60  
DB 1 MSFSLALSGVCTGLANVSKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60  
QY 61 VVASPTDNDPDIYTTWRDGLVLTIVDLFRNGDTSILSTIENYISAQAIVQGISNPSG 120  
DB 61 VVASPTDNDPDIYTTWRDGLVLTIVDLFRNGDTSILSTIENYISAQAIVQGISNPSG 120  
QY 121 DLSGAGLGEPEKFNVDERTAYTGSWGRPQRDPALRATMIFGQMLDNGTSTRTDIW 180  
DB 121 DLSGAGLGEPEKFNVDERTAYTGSWGRPQRDPALRATMIFGQMLDNGTSTRTDIW 180  
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFPIAVOHRALVEGSAFATAVGSSGSCWDSQ 240  
DB 171 PLVNDLSYVAQYNNQGYDLMEEVNGSSFPIAVOHRALVEGSAFATAVGSSGSCWDSQ 230  
QY 241 APEILCYLQSFMTGSFILANFDSRSRGKANTLGSIHFPPEACDDSTFQPCSPRALA 300  
DB 231 APEILCYLQSFMTGSFILANFDSRSRGKANTLGSIHFPPEACDDSTFQPCSPRALA 290  
QY 301 NHKEVDSFRSITYLNDGLDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360  
DB 291 NHKEVDSFRSITYLNDGLDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 350  
QY 361 KQGLLEVTDLDPFKALYSDATGTYSSSSSTYSIVDAKTFADGVSIIVETHAASNG 420  
DB 351 KQGLLEVTDLDPFKALYSDATGTYSSSSSTYSIVDAKTFADGVSIIVETHAASNG 410  
QY 421 SMSQYDKSDGEQLSARDLTWSYALLLTANNRRNSVVPASWGETSASSVPCTCATSAIG 480  
DB 411 SMSQYDKSDGEQLSARDLTWSYALLLTANNRRNSVVPASWGETSASSVPCTCATSAIG 470  
QY 481 TYSSVTVTWSRISYATGTTTATPTGSGTSTSKTATASKTSTTSS 530  
DB 471 TYSSVTVTWSRISYATGTTTATPTGSGTSTSKTATASKTSTTSS 520  
RESULT 11  
AAB15176  
ID AAB15176 standard; protein; 616 AA.  
AC AAB15176;  
XX  
DT 11-DEC-2000 (first entry)  
XX  
DE Aspergillus awamori glucoamylase.  
XX  
KW Glucoamylase; enzyme; carbohydrase; glucose;  
KM 1,4-alpha-D-glucan glucohydrolase.  
XX  
OS Aspergillus awamori.  
XX  
PN MO200043504-A1.  
XX  
PD 27-JUL-2000.  
XX  
PF 10-JAN-2000; 2000MO-US000532.  
XX  
PR 22-JAN-1999; 99US-002336063.  
XX  
PA (IOWA ) UNIV IOWA STATE RES POUND INC.  
XX  
PI Allen MJ, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatko R;  
FI Ford C;

XX WPI; 2000-514725/46.  
DR  
XX Fungal glucoamylase for selective production of glucose rather than alpha  
PT -1,6 linked disaccharide isomaltose, has mutation pair Asn20Cys coupled  
PT with Ala27Cys forming disulfide bond between the two stabilizing members.  
XX  
PS Disclosure; Page 152-153; 160pp; English.  
XX  
CC The present sequence is glucoamylase (1,4-alpha-D-glucan glucohydrolase;  
CC E.C. 3.2.1.3). This enzyme is a carbohydrase, and cleaves D-glucose from  
CC the nonreducing ends of maltotrioligosaccharides, attacking alpha-(1,4)-  
CC and alpha-(1,6)-glucosidic bonds. The present invention relates to mutant  
CC glucoamylases (see AAB15178-B15184), which have increased  
CC thermostability, increased pH optimum and reduced isomaltose formation.  
CC The mutant proteins are useful for the selective production of glucose  
CC rather than alpha-1,6 linked disaccharide isomaltose  
XX  
SQ Sequence 616 AA;  
Query Match 94.8%; Score 2628; DB 3; Length 616;  
Best Local Similarity 99.6%; Pred. No. 6,5e-203;  
Matches 504; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 25 ATLDSWLSNEATVARTAILNNIGADGAWSGADSGIYVVASPTDNDPDIYTTWRDGLVLT 84  
DB 1 ATLDSWLSNEATVARTAILNNIGADGAWSGADSGIYVVASPTDNDPDIYTTWRDGLVLT 60  
QY 85 KTLVLDLFRNGDTSILSTIENYISAQAIVQGISNPSGDLSSGAGLGEPEKFNVDERTAYTGSW 144  
DB 61 KTLVLDLFRNGDTSILSTIENYISAQAIVQGISNPSGDLSSGAGLGEPEKFNVDERTAYTGSW 120  
QY 145 GRPQRDPALRATAMIFGQMLDNGTSTRTDIWPLVNDLSYVAQYNNQGYDLMEE 204  
DB 121 GRPQRDPALRATAMIFGQMLDNGTSTRTDIWPLVNDLSYVAQYNNQGYDLMEE 180  
QY 205 VNGSSFPIAVOHRALVEGSAFATAVGSSGSCWDSQAPILCYLQSFMTGSFILANDSS 264  
DB 181 VNGSSFPIAVOHRALVEGSAFATAVGSSGSCWDSQAPILCYLQSFMTGSFILANDSS 240  
QY 265 RSGKANTLGSIHFPPEACDDSTFQPCSPRALAHKEVDSFRSITYLNDGLDSEEA 324  
DB 241 RSGKANTLGSIHFPPEACDDSTFQPCSPRALAHKEVDSFRSITYLNDGLDSEEA 300  
QY 325 VAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMDKQGLLEVTDSLDPFKALYSDAAT 384  
DB 301 VAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMDKQGLLEVTDSLDPFKALYSDAAT 360  
QY 385 GTYSSSSSTYSIVDAKTFADGVSIIVETHAASNGMSQYDKSDGEQLSARDLTWSYA 444  
DB 361 GTYSSSSSTYSIVDAKTFADGVSIIVETHAASNGMSQYDKSDGEQLSARDLTWSYA 420  
QY 445 ALLTANNRRNSVVPASWGETSASSVPCTCATSAIGTYSSVTVTWSRISYATGTTTAT 504  
DB 421 ALLTANNRRNSVVPASWGETSASSVPCTCATSAIGTYSSVTVTWSRISYATGTTTAT 480  
QY 505 PTGSGSYTSTSKTATASKTSTTSS 530  
DB 481 PTGSGSYTSTSKTATASKTSTTSS 506  
RESULT 12  
AAM55979  
ID AAM55979 standard; protein; 616 AA.  
AC AAM55979;  
XX  
DT 27-JUL-1998 (first entry)  
XX  
DE Aspergillus awamori glucoamylase mutant S411A.  
XX  
KW Aspergillus awamori; glucoamylase; Aspergillus sp. mutant; fungal; food;  
KM fructose; corn; sweetener; 1,4-alpha-D-glucan glucohydrolase;

KW genetic engineering.  
XX Synthetic.  
OS Aspergillus awamori.  
XX WO9803639-A1.  
XX PN  
XX PD 29-JAN-1998.  
XX PF 24-JUL-1997; 97WO-US012983.  
XX PR 24-JUL-1996; 96US-0022578P.  
XX PR 02-AUG-1996; 96US-0023077P.  
XX PA (IOWA ) UNIV IOWA STATE RES FOUND INC.  
XX PI Allen M, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatko R,  
PI Ford C;  
XX WPI; 1998-120764/11.  
XX DR  
XX PT Genetically engineered fungal glucosylase - useful in, e.g. food  
XX PT industry for production of high fructose corn sweeteners.  
XX PS Claim 10; Page: 97pp; English.  
XX CC The present sequence represents a specifically claimed mutant  
XX CC glucosylase from Aspergillus awamori (1,4-alpha-D-glucan  
XX CC glucosylase). The present invention describes fungal glucosylases  
XX CC (FG) comprising: a mutation pair Asn20Cys coupled with Ala27Cys forming a  
XX CC disulfide bond between the 2 members of the pair; and a 311-314loop or  
XX CC Ser11Ala mutation. FG can be used in industry for the production of high  
XX CC fructose corn sweeteners, while the glucose produced by glucosylase can  
XX CC be crystallised or used in fermentation to produce organic products, e.g.  
XX CC citric acid, ascorbic acid, lysine, glutamic acid or ethanol for  
XX CC beverages and fuel. The mutations provide increased thermal stability,  
XX CC reduced isomaltose formation and increased pH optimum. N.B. The present  
XX CC sequence is not given in the specification but is derived from SEQ ID  
XX CC NO.1 as stated in the claim  
SQ Sequence 616 AA;  
Query Match 94.7%; Score 2625; DB 2; Length 616;  
Best Local Similarity 99.4%; Pred. No. 1,1e-202; Indels 0; Gaps 0;  
Matches 503; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 25 ATLDWSLNEATVARTAILNINIGADGAWVSGADSGIVVASTDNDPDTFTWTRDSGLVL 84  
DB 1 ATLDWSLNEATVARTAILNINIGADGAWVSGADSGIVVASTDNDPDTFTWTRDSGLVL 60  
QY 85 KTLVULFRNGDTSLSTIENYISAQAIYVGISNPGDLSGGLGEPKENVDETATYTGSM 144  
DB 61 KTLVULFRNGDTSLSTIENYISAQAIYVGISNPGDLSGGLGEPKENVDETATYTGSM 120  
QY 145 GRPORDGALRATAMIGGOMLNDGYSTATDIWPLVRNLSVAOVNMGTDIWE 204  
DB 121 GRPORDGALRATAMIGGOMLNDGYSTATDIWPLVRNLSVAOVNMGTDIWE 180  
QY 205 VNGSSFFTIAYQHRALVEGSAFATVAGSSCSWCDQAPEILCYLOSFWTGSITLANPDS 264  
DB 181 VNGSSFFTIAYQHRALVEGSAFATVAGSSCSWCDQAPEILCYLOSFWTGSITLANPDS 240  
QY 265 RSGKANTLLGSIHFDPDEAACDDSTFQPCSPRALANHEVVDSTFRSYITLNDGSDSEA 324  
DB 241 RSGKANTLLGSIHFDPDEAACDDSTFQPCSPRALANHEVVDSTFRSYITLNDGSDSEA 300  
QY 325 VAVGRYPEDTYNGNPMFLCTLAABEOLYDALYQMDKQSLFVTVSILDFKALYSDAT 384  
DB 301 VAVGRYPEDTYNGNPMFLCTLAABEOLYDALYQMDKQSLFVTVSILDFKALYSDAT 360  
QY 385 GTYSSSSSTYSISVDAVKTFFADGPFVSVETHAASNGSSEQDYKSDGEOLASRDLTWSYA 444  
DB 361 GTYSSSSSTYSISVDAVKTFFADGPFVSVETHAASNGSSEQDYKSDGEOLASRDLTWSYA 420

QY 445 ALLTANNRNSVVPASWGETSASSVPGCAATSAIGTSSVTVTSPSTIVATGGTTTAT 504  
DB 421 ALLTANNRNSVVPASWGETSASSVPGCAATSAIGTSSVTVTSPSTIVATGGTTTAT 480  
QY 505 PTGSGSVTSTSKTATATASKTSTTSS 530  
DB 481 PTGSGSVTSTSKTATATASKTSTTSS 506

## RESULT 13

AAB15180  
ID AAB15180 standard; protein; 616 AA.

XX AAB15180;

XX 11-DEC-2000 (first entry)

XX Aspergillus awamori mutant glucosylase S41A substitution.

XX Glucosylase; enzyme; carbohydrase; glucose;  
XX 1,4-alpha-D-glucan glucosylase; muten; mutation.

XX Aspergillus awamori.

OS Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 411 /note="Wild-type Ser substituted by Ala"

XX WO20043504-A1.

XX 27-JUL-2000.

XX 10-JAN-2000; 2000WO-US000532.

XX 22-JAN-1999; 99US-00236063.

XX (IOWA ) UNIV IOWA STATE RES FOUND INC.

XX Allen MJ, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatko R,  
XX Ford C;

XX WPI; 2000-514725/46.

XX Fungal glucosylase for selective production of glucose rather than alpha  
XX -1,6 linked disaccharide isomaltose, has mutation pair Asn20Cys coupled  
XX with Ala27Cys forming disulfide bond between the two stabilizing members.

XX Claim 10; Page: 160pp; English.

XX Glucosylase (1,4-alpha-D-glucan glucosylase; E.C. 3.2.1.3) is a  
XX carbohydrase. This enzyme cleaves D-glucose from the nonreducing ends of  
XX maltooligosaccharides, attacking alpha-(1,4)-, and alpha-(1,6)-glucosidic  
XX bonds. The present invention relates to mutant glucosylases, which have  
XX increased thermostability, increased pH optimum and reduced isomaltose  
XX formation. The mutant proteins are useful for the selective production of  
XX glucose rather than alpha-1,6 linked disaccharide isomaltose. The present  
XX sequence is one such mutant enzyme. Note: The present sequence is not  
XX shown in the specification but is derived from the Aspergillus awamori  
XX wild-type glucosylase sequence given in pages 152-153 of the sequence  
XX listing (SEQ ID 1)

XX Sequence 616 AA;

Query Match 94.7%; Score 2625; DB 3; Length 616;  
Best Local Similarity 99.4%; Pred. No. 1,1e-202; Indels 0; Gaps 0;  
Matches 503; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 25 ATLDWSLNEATVARTAILNINIGADGAWVSGADSGIVVASTDNDPDTFTWTRDSGLVL 84  
DB 1 ATLDWSLNEATVARTAILNINIGADGAWVSGADSGIVVASTDNDPDTFTWTRDSGLVL 60

435  
Not claimed

QY	85	KLVLDFPENGDTSLSTLENIYSALIVOGISNPSGDLSSGAGIAGBKPFVDETAATYGS*	144
Db	61	KLVLDFRNGDTSLSLSTLENIYSALQALVQGISNPSGDLSSGAGIAGBKPFVDETAATYGS*	120
QY	145	GRPQDGPALRATAMIGFGOMLLDNGYSTATDIIWPLVLRNDLSVYAQYWNQGYDLME*	204
Db	121	GRPQDGPALRATAMIGFGOMLLDNGYSTATDIIWPLVLRNDLSVYAQYWNQGYDLME*	180
QY	205	VNGSSFPIITAVQHALVEGSAFATAVAGSSCSMCDQAPBILCYLQSFPMWGSFI LANPDS*	264
Db	181	VNGSSFPIITAVQHALVEGSAFATAVAGSSCSMCDQAPBILCYLQSFPMWGSFI LANPDS*	240
QY	265	RSKGANTLIGSIHTFDPPEAACDDSTFOCBSPRALNHEKVDSFRSITILNDGLSDSEA*	324
Db	241	RSKGANTLIGSIHTFDPPEAACDDSTFOCBSPRALNHEKVDSFRSITILNDGLSDSEA*	300
QY	325	VAVGRIPEDDTYNNGNPMFLCTLLAAAEQLYDALYQMDKQSLSEVTDVSLDFPKALYSDAAT*	384
Db	301	VAVGRIPEDDTYNNGNPMFLCTLLAAAEQLYDALYQMDKQSLSEVTDVSLDFPKALYSDAAT*	360
QY	385	GTYSSSSSTYSISIVDAVKTFADGFPYSIVETHAASNGSMSEQYKXSGEOLSAADLTWSYA*	444
Db	361	GTYSSSSSTYSISIVDAVKTFADGFPYSIVETHAASNGSMSEQYKXSGEOLSAADLTWSYA*	420
QY	445	ALLTANNRRNSVVPASWGETSASVPGTCAATSAIGTYSVVTWSPISIVATGTTTAT*	504
Db	421	ALLTANNRRNSVVPASWGETSASVPGTCAATSAIGTYSVVTWSPISIVATGTTTAT*	480
QY	505	PTGSGSVTSTKTTATASTKSTTTS*	530
Db	481	PTGSGSVTSTKTTATASTKSTTTS*	506

RESULT 14	
AAW55976	
ID	AAW55976 standard; protein; 616 AA.
XX	
AC	
XX	AAW55976;
XX	
DT	27-JUL-1998 (first entry)
XX	
DE	Aspergillus awamori glucoamylase mutant N20C, A27C.
XX	
KW	Aspergillus awamori; glucoamylase; Aspergillus sp; mutant; fungal; food;
KW	fructose; corn; sweetener; 1,4-alpha-D-glucan glucosyltransferase;
KW	genetic engineering.
XX	
OS	Synthetic.
OS	Aspergillus awamori.
XX	
FH	Key
FT	Disulfide-bond 20.27
XX	
PN	WO9803639-A1.
XX	
PD	29-JAN-1998.
XX	
XX	24-JUL-1997; 97WO-US012983.
XX	
PP	24-JUL-1996; 96US-0032578P.
PR	02-AUG-1996; 96US-0023077P.
XX	
XX	(IOWA ) UNIV IOWA STATE RES FOUND INC.
XX	
PI	Allen M, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatko R;
PI	Ford C;
XX	
DR	WPI; 1998-120764/11.
XX	
PT	Genetically engineered fungal glucoamylase - useful in, e.g. food
PT	industry for production of high fructose corn sweeteners.
XX	
P9	Claim 1; Page; 97pp; English.

XX The present sequence represents a specifically claimed mutant  
CC glycoamylase from *Ampelillus awamori* (1,4- $\alpha$ -D-glucan  
CC glycohydrolase). The present invention describes fungal glycoamylases  
CC (FG) comprising: a mutation pair Asn20Cys coupled with Ala27Cys forming a  
CC disulphide bond between the 2 members of the pair; and a 311-314Loop or  
CC Ser11Aa mutation. FG can be used in industry for the production of high  
CC fructose corn sweetener, while the glucose produced by glycoamylase can  
CC be crystallised or used in fermentation to produce organic products, e.g.  
CC citric acid, ascorbic acid, lysine, glutamic acid or ethanol for  
CC beverages and fuel. The mutations provide increased thermal stability,  
CC reduced isomaltose formation and increased pH optimum. N.B. The present  
CC sequence is not given in the specification but is derived from SEQ ID  
CC No.1 as stated in the claim  
XX  
XX Sequence 616 AA;  
SQ

Query Match	94.4%;	Score 2615;	DB 2;	Length 616;
Best Local Similarly	99.2%;	Pred. No. 7.3e-202;		
Matches 502; Conservative	1;	Mismatches 3;	Indels 0;	Gaps 0;

QY	25	ATLDSWISNEMTAVRTALINNI	IGD	GA	VSG	DSG	SI	V	AS	STDP	ND	P	F	Y	T	R	S	G	I	V	L	84																																		
Db	1	ATLDSWISNEMTAVRTALIN	IG	AD	GA	CV	SG	AD	SG	SI	V	AS	STDP	ND	P	F	Y	T	R	S	G	I	V	L	60																															
QY	85	KTLYDLFPNCGTSL	SL	ST	EN	Y	IS	SA	O	I	Y	Q	I	S	N	P	S	G	D	L	S	S	G	A	G	L	E	P	K	N	N	D	E	T	A	I	T	S	G	M	144															
Db	61	KTLYDLFPNCGTSL	SL	ST	EN	Y	IS	SA	O	I	Y	Q	I	S	N	P	S	G	D	L	S	S	G	A	G	L	E	P	K	N	N	D	E	T	A	I	T	S	G	M	120															
QY	145	GRPODGEAL	B	A	T	A	M	I	G	F	Q	W	L	L	D	N	G	Y	S	T	A	T	D	I	W	P	L	V	R	N	D	L	S	V	A	Y	A	N	N	O	T	G	D	L	M	E	204									
Db	121	GRPODGEAL	B	A	T	A	M	I	G	F	Q	W	L	L	D	N	G	Y	S	T	A	T	D	I	W	P	L	V	R	N	D	L	S	V	A	Y	A	N	N	O	T	G	D	L	M	E	180									
QY	205	VNGSS	F	F	T	I	A	V	O	H	R	A	L	V	E	G	S	A	P	A	T	A	V	S	G	S	C	S	W	C	D	S	O	A	P	E	L	L	C	Y	O	S	F	M	T	G	S	F	I	L	N	F	D	S	264	
Db	181	VNGSS	F	F	T	I	A	V	O	H	R	A	L	V	E	G	S	A	P	A	T	A	V	S	G	S	C	S	W	C	D	S	O	A	P	E	L	L	C	Y	O	S	F	M	T	G	S	F	I	L	N	F	D	S	240	
QY	265	RS	G	D	A	N	T	L	L	S	I	H	T	F	P	E	A	C	D	S	T	P	O	C	P	R	A	L	N	K	E	V	D	S	F	R	S	I	Y	T	L	N	D	G	L	S	D	S	E	A	324					
Db	241	RS	G	D	A	N	T	L	L	S	I	H	T	F	P	E	A	C	D	S	T	P	O	C	P	R	A	L	N	K	E	V	D	S	F	R	S	I	Y	T	L	N	D	G	L	S	D	S	E	A	300					
QY	325	V	A	V	A	R	P	E	D	T	Y	N	G	N	P	W	F	L	C	T	L	A	A	B	O	U	Y	D	A	L	Y	Q	N	D	K	O	G	S	L	E	V	T	D	S	L	D	F	F	K	A	L	S	D	A	T	384
Db	301	V	A	V	A	R	P	E	D	T	Y	N	G	N	P	W	F	L	C	T	L	A	A	B	O	U	Y	D	A	L	Y	Q	N	D	K	O	G	S	L	E	V	T	D	S	L	D	F	F	K	A	L	S	D	A	T	360
QY	385	G	T	Y	S	S	S	S	S	S	Y	S	I	D	A	V	K	T	P	A	D	F	P	S	I	Y	E	T	H	A	S	G	S	M	S	E	O	Y	D	S	D	G	O	L	S	A	R	D	L	T	W	S	A	444		
Db	361	G	T	Y	S	S	S	S	S	S	Y	S	I	D	A	V	K	T	P	A	D	F	P	S	I	Y	E	T	H	A	S	G	S	M	S	E	O	Y	D	S	D	G	O	L	S	A	R	D	L	T	W	S	A	420		
QY	445	A	L	L	T	A	N	N	R	N	S	V	V	P	A	S	M	G	E	T	S	A	S	S	V	P	G	C	A	T	S	A	I	G	T	S	S	V	T	S	M	P	S	I	V	A	N	G	T	T	T	A	504			
Db	421	A	L	L	T	A	N	N	R	N	S	V	V	P	A	S	M	G	E	T	S	A	S	S	V	P	G	C	A	T	S	A	I																							

RESULT 15  
 AAB15178  
 ID AAB15178 standard; protein; 616 AA.  
 XX  
 AC AAB15178;  
 XX  
 DT 11-DEC-2000 (first entry)  
 XX  
 DE Aspergillus awamori mutant glucocamylase N20C/A27C substitution.  
 XX  
 XX  
 KW Glucoamylase; enzyme; carbohydrase; glucose;  
 KM 1,4-alpha-D-glucan glucohydrolase; mutein; mutation.  
 XX  
 OS Aspergillus awamori.  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH Disulfide-bond 20..27  
 FT Misc-difference 20  
 FT Misc-difference 27 /note= "wild-type Asn substituted by Cys"  
 FT Misc-difference 27 /note= "wild-type Ala substituted by Cys"  
 XX MO200043504-A1.  
 XX 27-JUL-2000.  
 XX 10-JAN-2000; 2000WO-US000532.  
 XX 22-JAN-1999; 99US-00236063.  
 XX (IOWA ) UNIV IOWA STATE RES FOUND INC.  
 XX Allen MJ, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatko R,  
 PI Ford C;  
 XX WPI; 2000-514725/46.  
 XX  
 XX Fungal glucosylase for selective production of glucose rather than alpha  
 PT -1,6 linked disaccharide isomaltose, has mutation pair Asn20Cys coupled  
 PT with Ala27Cys forming disulfide bond between the two stabilizing members.  
 XX  
 XX Claim 1; Page; 160pp; English.  
 XX  
 XX Glucosylase (1,4-alpha-D-glucan glucosylase; E.C. 3.2.1.3) is a  
 CC carbohydriase. This enzyme cleaves D-glucose from the nonreducing ends of  
 CC maltooligosaccharides, attacking alpha-(1,4)-, and alpha-(1,6)-glucosidic  
 CC bonds. The present invention relates to mutant glucosylases, which have  
 CC increased thermostability, increased pH optimum and reduced isomaltose  
 CC formation. The mutant proteins are useful for the selective production of  
 CC glucose rather than alpha-1,6 linked disaccharide isomaltose. The present  
 CC sequence is one such mutant enzyme. Note: The present sequence is not  
 CC shown in the specification but is derived from the Aspergillus awamori  
 CC wild-type glucosylase sequence given in pages 152-153 of the sequence  
 CC listing (Seq ID 1)  
 CC  
 XX Sequence 616 AA;  
 SQ

Query Match 94.4%; Score 2615; DB 3; Length 616;  
 Best Local Similarity 99.2%; Pred. No. 7.3e-202;  
 Matches 502; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 25 ATLDMSLNENAVARTAILNINIGADGAWSGADSGIVVAPSTNDNDYFYTWTRDSGLVL 84  
 DB 1 ATLDMSLNENAVARTAILNINIGADGAWSGADSGIVVAPSTNDNDYFYTWTRDSGLVL 60  
 QY 85 KTLVDLFRNGDSTLSLTIBENYISAOAIYVQISNPSGDLSSGAGLGEPEKFNDEYATYGSW 144  
 DB 61 KTLVDLFRNGDSTLSLTIBENYISAOAIYVQISNPSGDLSSGAGLGEPEKFNDEYATYGSW 120  
 QY 145 GRPORDGPRALRTAMIGFGOMLNDNGYSTARDIWPPLVRNDLSYVAQVWNOTGYDLWE 204  
 DB 121 GRPORDGPRALRTAMIGFGOMLNDNGYSTARDIWPPLVRNDLSYVAQVWNOTGYDLWE 180  
 QY 205 VNGSSFFTIAYQHRALVEGSAFATVAGSSCSWCDSDQAPETLCYLOSFWTGSFILANFDS 264  
 DB 181 VNGSSFFTIAYQHRALVEGSAFATVAGSSCSWCDSDQAPETLCYLOSFWTGSFILANFDS 240  
 QY 265 RSGKDANTILGSIHTFDPBAACDSTFQCSPRALANHKEVVDSPRSIYTLNDGLSDSEA 324  
 DB 241 RSGKDANTILGSIHTFDPBAACDSTFQCSPRALANHKEVVDSPRSIYTLNDGLSDSEA 300  
 QY 325 VAVGRYPEDTYNGNPMFLCTLAABQLYDALYQWDKQSLFETDVSLEDFKALYSDAAT 384  
 DB 301 VAVGRYPEDTYNGNPMFLCTLAABQLYDALYQWDKQSLFETDVSLEDFKALYSDAAT 360  
 QY 385 GTYSSSSSTYSIVAVKTFADGPFVSIETHAASNGSMSEQYDKSDGEQLSARDLTWSYA 444

DB 361 GTYSSSSSTYSIVAVKTFADGPFVSIETHAASNGSMSEQYDKSDGEQLSARDLTWSYA 420  
 QY 445 ALLTANNRRNSVPAKMGETSASVPGCAATSAIGTYSVTYTSWPSIVATGCTTTAT 504  
 DB 421 ALLTANNRRNSVPAKMGETSASVPGCAATSAIGTYSVTYTSWPSIVATGCTTTAT 480  
 QY 505 PTGSGSVTSTSKTATATASKTSTTTS 530  
 DB 481 PTGSGSVTSTSKTATATASKTSTTTS 506

Search completed: June 17, 2004, 17:45:59  
 Job time : 62 secs

**HIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:47:36 ; Search time 50 Seconds  
(without alignments)  
3015.110 Million cell updates/sec

Title: US-10-038-723-2  
Perfect score: 2771  
Sequence: 1 MSFSLSLSGVCTGLANV.....SKTTATSKTSTTTTSGMSL 534

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues  
Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2771	100.0	534	US-09-908-395-2	Sequence 2, Appli
2	2771	100.0	534	US-09-821-616-9	Sequence 9, Appli
3	2771	100.0	534	US-10-038-723-2	Sequence 2, Appli
4	2771	100.0	534	US-10-421-586-2	Sequence 2, Appli
5	2771	100.0	534	US-10-421-586-3	Sequence 3, Appli
6	2742	99.0	640	US-10-038-723-13	Sequence 13, Appli
7	2742	99.0	640	US-10-421-586-14	Sequence 14, Appli
8	2735	98.7	743	US-10-418-836-9	Sequence 9, Appli
9	2735	98.7	763	US-10-418-836-13	Sequence 13, Appli
10	2735	98.7	979	US-10-418-836-10	Sequence 10, Appli
11	2735	98.7	979	US-10-418-836-16	Sequence 16, Appli
12	2720.5	98.2	738	US-10-418-836-19	Sequence 19, Appli
13	2716	98.0	741	US-10-418-836-30	Sequence 30, Appli
14	2715.5	98.0	972	US-10-418-836-18	Sequence 38, Appli
15	2715	98.0	739	US-10-418-836-26	Sequence 26, Appli

16	2715	98.0	739	US-10-418-836-27	Sequence 27, Appli
17	2714.5	98.0	740	US-10-418-836-28	Sequence 28, Appli
18	2714.5	98.0	740	US-10-418-836-34	Sequence 34, Appli
19	2714.5	98.0	742	US-10-418-836-29	Sequence 29, Appli
20	2714	97.9	743	US-10-418-836-31	Sequence 31, Appli
21	2714	97.9	743	US-10-418-836-35	Sequence 35, Appli
22	2714	97.9	975	US-10-418-836-39	Sequence 39, Appli
23	2523	91.1	1095	US-10-228-063-45	Sequence 45, Appli
24	1850.5	66.8	631	US-10-213-990-54	Sequence 54, Appli
25	1689	61.0	618	US-09-821-616-74	Sequence 34, Appli
26	1674.5	60.4	591	US-09-821-616-7	Sequence 7, Appli
27	1548.5	55.9	647	US-10-369-493-1109	Sequence 3109, Appli
28	1425.5	51.4	581	US-10-281-673-3	Sequence 3, Appli
29	1408.5	50.8	704	US-10-213-990-21	Sequence 21, Appli
30	1380.5	49.8	620	US-10-213-990-33	Sequence 33, Appli
31	1250	45.1	441	US-10-369-493-12402	Sequence 12402, A
32	1132	40.9	432	US-10-369-493-12533	Sequence 12533, A
33	781.5	28.2	450	US-10-369-493-2490	Sequence 2490, Ap
34	725	26.2	579	US-10-228-063-49	Sequence 49, Appli
35	597	21.5	549	US-10-369-493-22096	Sequence 22096, A
36	164.5	5.9	2344	US-09-815-242-12713	Sequence 12713, A
37	160	5.8	35	US-09-821-616-4	Sequence 4, Appli
38	159.5	5.8	2283	US-10-172-502-4	Sequence 4, Appli
39	159	5.7	2271	US-10-282-122A-43924	Sequence 43924, A
40	158	5.7	1283	US-10-369-493-22616	Sequence 22616, A
41	155.5	5.6	1621	US-10-185-990-11	Sequence 11, Appli
42	155.5	5.6	1626	US-10-185-990-11	Sequence 11, Appli
43	154.5	5.6	1031	US-09-815-242-10932	Sequence 10932, A
44	152.5	5.5	2117	US-10-120-801-63	Sequence 63, Appli
45	152	5.5	45	US-10-228-063-53	Sequence 53, Appli

## ALIGNMENTS

RESULT 1  
US-09-908-395-2  
; Sequence 2, Application US/09908395  
; Patent No. US20020164723A1  
; GENERAL INFORMATION:  
; APPLICANT: liaw, Gai  
; APPLICANT: Pedersen, Sven  
; APPLICANT: Hendriksen, Sven  
; TITLE OF INVENTION: A Method of Producing Saccharide  
; TITLE OF INVENTION: Preparations  
; FILE REFERENCE: 5318.200-US  
; CURRENT APPLICATION NUMBER: US/09/908.395  
; CURRENT FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US/09/198.672  
; PRIOR FILING DATE: PRIOR FILING DATE: 1998-11-23  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Aspergillus Niger  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(24)  
US-09-908-395-2

Query Match 100.0%; Score 2771; DB 9; Length 534;  
Best Local Similarity 100.0%; Pred. No. 2.7e-237;  
Matches 534; Conservative 0; Mismatches 0; Gaps 0;

QY 1 MSFSLSLSGVCTGLANVSKATLDSMTSNEATVARTIILNNGAGAVSGADSGI 60  
DB 1 MSFSLSLSGVCTGLANVSKATLDSMTSNEATVARTIILNNGAGAVSGADSGI 60  
QY 61 VVASPSTNDPYPYFTWTRDSGLVLTVDLFRNEDTSLSTIENYISAQAIVOGISNPSG 120  
DB 61 VVASPSTNDPYPYFTWTRDSGLVLTVDLFRNEDTSLSTIENYISAQAIVOGISNPSG 120

Post-processed  
-a-p-r-i-c-a-m

```
QY 121 DSSGAGLGEPKRVNDETAATYTGSGWRPQDGPALRAIAMIQFGQWLLDNGYTSTATDIW 180
D 121 DSSGAGLGEPKRVNDETAATYTGSGWRPQDGPALRAIAMIQFGQWLLDNGYTSTATDIW 180
QY 181 PLVBNDSLVAQVWNONQGYLMEEVNCSPTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
D 181 PLVBNDSLVAQVWNONQGYLMEEVNCSPTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
QY 241 APEILCYLQSFMTGSLFIANFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPALA 300
D 241 APEILCYLQSFMTGSLFIANFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPALA 300
QY 301 NHKEVDSFRSIIYTLNDGLSDSEAVAGRYPEDTYNGNPMFLCTLLAAEQLYDALYQMD 360
D 301 NHKEVDSFRSIIYTLNDGLSDSEAVAGRYPEDTYNGNPMFLCTLLAAEQLYDALYQMD 360
QY 361 KQGSLEVTDVSLDPEFKALYSDAATGTYSSSSSTYSIVDAVKTPADGFSIVETHAASNG 420
D 361 KQGSLEVTDVSLDPEFKALYSDAATGTYSSSSSTYSIVDAVKTPADGFSIVETHAASNG 420
QY 421 SMSEQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
D 421 SMSEQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
QY 481 TYSSVTVTSWPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
D 481 TYSSVTVTSWPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

```
RESULT 2
US-09-821-616-9
; Sequence 9, Application US/09821616
; Publication No. US20030027290A1
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Rudy
; APPLICANT: Lembeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279, 200-US
; CURRENT APPLICATION NUMBER: US/09/821, 616
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/199, 290
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1557/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 0925/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/070, 746
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/094, 344
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/979, 673
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/107, 657
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-821-616-9
```

```
Query Match 100.0%; Score 2771; DB 10; Length 534;
Best Local Similarity 100.0%; Pred. No. 2,7e-237;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFSLALSLGLVCTGLANVYSKRAITLDSWLSNEATYATKTIILNNGADGAWGSGANSGI 60
D 1 MSFSLALSLGLVCTGLANVYSKRAITLDSWLSNEATYATKTIILNNGADGAWGSGANSGI 60
```

```
QY 61 VVASPTDNPDYFYTWTRDSGLVLTLYDLFRNGDTSLSTIENYISAQAIYQGISNPSG 120
D 61 VVASPTDNPDYFYTWTRDSGLVLTLYDLFRNGDTSLSTIENYISAQAIYQGISNPSG 120
QY 121 DSSGAGLGEPKRVNDETAATYTGSGWRPQDGPALRAIAMIQFGQWLLDNGYTSTATDIW 180
D 121 DSSGAGLGEPKRVNDETAATYTGSGWRPQDGPALRAIAMIQFGQWLLDNGYTSTATDIW 180
QY 181 PLVBNDSLVAQVWNONQGYLMEEVNCSPTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
D 181 PLVBNDSLVAQVWNONQGYLMEEVNCSPTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
QY 241 APEILCYLQSFMTGSLFIANFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPALA 300
D 241 APEILCYLQSFMTGSLFIANFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPALA 300
QY 301 NHKEVDSFRSIIYTLNDGLSDSEAVAGRYPEDTYNGNPMFLCTLLAAEQLYDALYQMD 360
D 301 NHKEVDSFRSIIYTLNDGLSDSEAVAGRYPEDTYNGNPMFLCTLLAAEQLYDALYQMD 360
QY 361 KQGSLEVTDVSLDPEFKALYSDAATGTYSSSSSTYSIVDAVKTPADGFSIVETHAASNG 420
D 361 KQGSLEVTDVSLDPEFKALYSDAATGTYSSSSSTYSIVDAVKTPADGFSIVETHAASNG 420
QY 421 SMSEQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
D 421 SMSEQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
QY 481 TYSSVTVTSWPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
D 481 TYSSVTVTSWPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

```
RESULT 3
US-10-038-723-2
; Sequence 2, Application US/10038723
; Publication No. US20030032163A1
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne Roenfeldt
; APPLICANT: Svendsen, Allan
; APPLICANT: Pedersen, Henrik
; APPLICANT: Vinth, Jesper
; APPLICANT: Hendiksen, Hanne Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OF INVENTION: Glucoamylase Variants
; FILE REFERENCE: 5636, 200-US
; CURRENT APPLICATION NUMBER: US/10/038, 723
; CURRENT FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/351, 814
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 00937
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 01667
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093, 528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/115, 545
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-10-038-723-2
```

```
Query Match 100.0%; Score 2771; DB 14; Length 534;
Best Local Similarity 100.0%; Pred. No. 2,7e-237;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



QY	1	MSFRSLAASGLVCTGLANAVISGRATLDSWLSNEAVARPAIANNIGADAWYSGADSGI	60
Db	1	MSFRSLAASGLVCTGLANAVISGRATLDSWLSNEAVARPAIANNIGADAWYSGADSGI	60
QY	61	VVASPSTNDPBYFYTWTRDSGLVAKTLVDLFRNGDTSLLSTIENYISAQAIVOGISNPSG	120
Db	61	VVASPSTNDPBYFYTWTRDSGLVAKTLVDLFRNGDTSLLSTIENYISAQAIVOGISNPSG	120
QY	121	DLSGAGLGEPPENVDETAYTGSNGRPORGPALPAPAMGFEQWLNDNYSTANDIYW	180
Db	121	DLSGAGLGEPPENVDETAYTGSNGRPORGPALPAPAMGFEQWLNDNYSTANDIYW	180
QY	181	PLVRNDLSYVAQYWNQGYDLMEBVGNSSEFTIYVQHRALVEBSAPATAYAGSSCSWCDSQ	240
Db	181	PLVRNDLSYVAQYWNQGYDLMEBVGNSSEFTIYVQHRALVEBSAPATAYAGSSCSWCDSQ	240
QY	241	APBIIICYLOSFWTGSFILLNPFSSRBCKXANTLGSIHFFDPBEAACDDSTFFQCSPPALA	300
Db	241	APBIIICYLOSFWTGSFILLNPFSSRBCKXANTLGSIHFFDPBEAACDDSTFFQCSPPALA	300
QY	301	NHKEVVDSEFRSIYTLINDGJSDSEBAVAVGRIPEBPTYNNGNMFELCTIAAAEQLDALYQWD	360
Db	301	NHKEVVDSEFRSIYTLINDGJSDSEBAVAVGRIPEBPTYNNGNMFELCTIAAAEQLDALYQWD	360
QY	361	KOGSLEWTVSGLDFFPALYSDAATGTYSSSSSTYSSTIYDAVKTFPAQFUSIVETHAASNG	420
Db	361	KOGSLEWTVSGLDFFPALYSDAATGTYSSSSSTYSSTIYDAVKTFPAQFUSIVETHAASNG	420
QY	421	SMSBQYDKSDGEOQLSARDLTWYSYALLLTANNRNSVVPASMGESTSASVVGTCATSAIG	480
Db	421	SMSBQYDKSDGEOQLSARDLTWYSYALLLTANNRNSVVPASMGESTSASVVGTCATSAIG	480
QY	481	TYSSSTYTVSWBSIVATGCTTTTATPPGSSGSVYSTKTTATATSKTSTTTSGMGL	534
Db	481	TYSSSTYTVSWBSIVATGCTTTTATPPGSSGSVYSTKTTATATSKTSTTTSGMGL	534

Dolphins  
Yes

```

US-10-421-586-2
RESULT 4
: US-10-421-586-2
: Sequence 2, Application US/10421586
: Publication No. US20040002142A1
: GENERAL INFORMATION:
: APPLICANT: Melsen, Bjarne Ronfeldt
: APPLICANT: Svendsen, Allan
: APPLICANT: Pedersen, Henrik
: APPLICANT: Wind, Jesper
: APPLICANT: Hendriksen, Hanne Vang
: APPLICANT: Frandsen, Torben Peter
: TITLE OF INVENTION: Glucosylase Variants
: FILE REFERENCE: 5967, 210-US
: CURRENT APPLICATION NUMBER: US/10/421,586
: CURRENT FILING DATE: 2003-04-23
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 2
: LENGTH: 534
: TYPE: prt
: ORGANISM: Aspergillus niger
US-10-421-586-2

```

	Query Match	Similarity	Score	DB	Length
Best Local	534	100.0%	2771	21	534
Matches	534	Conservative	0	Mismatches	0
			Indels	0	Gaps
					0

Qy	121	LLSSAGJGEPKENVDETAAYGSGWRPORDEGALATAMIGGOMLLDNGYSTADIYW	180
Dp	121	DLSSAGJGEPKENVDETAAYGSGWRPORDEGALATAMIGGOMLLDNGYSTADIYW	180
Qy	181	PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFPTIAVOHRALVEGSAATAVSSCSWCDSQ	240
Db	181	PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFPTIAVOHRALVEGSAATAVSSCSWCDSQ	240
Qy	241	APELLICYIQSFWTGSFIILANFDSNSGKDANTLLSIRHPDEBAACDSDTPOCSPRALA	300
Dp	241	APELLICYIQSFWTGSFIILANFDSNSGKDANTLLSIRHPDEBAACDSDTPOCSPRALA	300
Qy	301	NHKEVYVDSFRSITLYINDGLSDSEAVAVGRYPEDTYNGNPNFLCTLAALAEOLYDALIYOMD	360
Dp	301	NHKEVYVDSFRSITLYINDGLSDSEAVAVGRYPEDTYNGNPNFLCTLAALAEOLYDALIYOMD	360
Qy	361	KQGSLEVTDVSLDPEFKALYSDPATGYSSSSSTYSIVDAVXTFADGFPVSIYETHAASNG	420
Dp	361	KQGSLEVTDVSLDPEFKALYSDPATGYSSSSSTYSIVDAVXTFADGFPVSIYETHAASNG	420
Qy	421	SMSBQYDKSDGEOQLSARDLWTYSYALLPLANNRNRNVVPASWQETASASSVPGTCAATSALIG	480
Dp	421	SMSBQYDKSDGEOQLSARDLWTYSYALLPLANNRNRNVVPASWQETASASSVPGTCAATSALIG	480
Qy	481	TYSSVTWYVWBPISYATGGTTTATPAGTSGSVTSTKTTATASTKSTTTTSGMSL	534
Dp	481	TYSSVTWYVWBPISYATGGTTTATPAGTSGSVTSTKTTATASTKSTTTTSGMSL	534

```

RESULT 5
US-10-421-586-3
; Sequence 3, Application US/10421586
; Publication No. US20040002142A1
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjørne Ronfeldt
; APPLICANT: Svendsen, Allan
; APPLICANT: Pedersen, Henrik
; APPLICANT: Vind, Jesper
; APPLICANT: Hendriksen, Haane Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OF INVENTION: Glucosylase Variants
; FILE REFERENCE: 5967,210-US
; CURRENT APPLICATION NUMBER: US/10/421,586
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: Signal
; LOCATION: (1)..(24)
; US-10-421-586-3

```

	Query Match	100.0%;	Score 2771;	DB 15;	Length 534;	
	Best Local Similarity	100.0%;	Pred. No. 2,7e+237;			
	Matches	534;	Conservative	0;	Mismatches	0;
					Indels	0;
					Gaps	0;
Qy	1	MSFSSLALSGLVCTGGLANVISKRAATLDSWLSNEATVARTAILINNTIGADGAMVSGADSGI	60			
Db	1	MSFSSLALSGLVCTGGLANVISKRAATLDSWLSNEATVARTAILINNTIGADGAMVSGADSGI	60			
Qy	61	VVASPSTDNPDYFYTTTRDSGLVVKTLVDLFRNGDTSLLSTIENYISAOAIVOGISNPSG	120			
Db	61	VVASPSTDNPDYFYTTTRDSGLVVKTLVDLFRNGDTSLLSTIENYISAOAIVOGISNPSG	120			
Qy	121	DLSSGAGLGEEKFVNDETAATYGSWGRPORDEPALRATATMIGFGWILDNGTSTATDILVM	180			
Db	121	DLSSGAGLGEEKFVNDETAATYGSWGRGORDEPALRATATMIGFGWILDNGTSTATDILVM	180			
Qy	181	PLVENDLSYVAQYVNOGTGYDIMEEVNGSSFTTIAVQHALVEGSAFATAVSSCSGCWCSQ	240			

```
Db      181 PLVNDLSYYAQYWNQGYDLMEBVNGSSPFTIAVQHRALVEGSAFATAVGSSCWCDSQ 240
Qy      241 APELICYLQSFMTGSLFIANFDSRSRSGKDNATLLGSIHTFDPBAACDSTFQPCSPRALA 300
Db      241 APELICYLQSFMTGSLFIANFDSRSRSGKDNATLLGSIHTFDPBAACDSTFQPCSPRALA 300
Qy      301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAEQLYDALYQMD 360
Db      301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAEQLYDALYQMD 360
Qy      361 KQSLLEVTVDSLDFPKALYSDAATGYSSSSSTYSIVDAVKTPADGFVSIIVETHAASNG 420
Db      361 KQSLLEVTVDSLDFPKALYSDAATGYSSSSSTYSIVDAVKTPADGFVSIIVETHAASNG 420
Qy      421 SMSEQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Db      421 SMSEQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Qy      481 TYSQVTVTSWPSIVATGTTTATPTGSGSVTSTSKTTATATASKTSTTRSGMSL 534
Db      481 TYSQVTVTSWPSIVATGTTTATPTGSGSVTSTSKTTATATASKTSTTRSGMSL 534
```

## RESULT 6

```
US-10-038-723-13
; Sequence 13, Application US/10038723
; Publication No. US20030032163A1
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne Roenfeldt
; APPLICANT: Svendsen, Allan
; APPLICANT: Pedersen, Henrik
; APPLICANT: Vind, Jesper
; APPLICANT: Hendriksen, Hanne Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OR INVENTION: Glucoamylase Variants
; FILE REFERENCE: 5636.200-US
; CURRENT APPLICATION NUMBER: US/10/038, 723
; PRIOR APPLICATION NUMBER: 09/351,814
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 00937
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 01667
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/115,545
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-12
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 13
; LENGTH: 640
; TYPE: PRT
; ORGANISM: ASPERGILLUS NIGER
US-10-038-723-13
```

```
Query Match      99.0%; Score 2742; DB 14; Length 640;
Best Local Similarity 99.6%; Pred. No. 1.4e-234;
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Db      1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Qy      61 VVASPSNDNDPYFYTWTRDSGLVLTLYDLFRNGDTSLSLTENYISAQAIVQGISNPSG 120
Db      61 VVASPSNDNDPYFYTWTRDSGLVLTLYDLFRNGDTSLSLTENYISAQAIVQGISNPSG 120
Qy      121 DLSSGAGLGEPPKFNVDETAYTGSWGRPQDGPALRATAMIGFGQWLNDNGYTSRATDIW 180
Db      121 DLSSGAGLGEPPKFNVDETAYTGSWGRPQDGPALRATAMIGFGQWLNDNGYTSRATDIW 180
Qy      181 PLVNDLSYYAQYWNQGYDLMEBVNGSSPFTIAVQHRALVEGSAFATAVGSSCWCDSQ 240
```

```
Db      181 PLVNDLSYYAQYWNQGYDLMEBVNGSSPFTIAVQHRALVEGSAFATAVGSSCWCDSQ 240
Qy      241 APELICYLQSFMTGSLFIANFDSRSRSGKDNATLLGSIHTFDPBAACDSTFQPCSPRALA 300
Db      241 APELICYLQSFMTGSLFIANFDSRSRSGKDNATLLGSIHTFDPBAACDSTFQPCSPRALA 300
Qy      301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAEQLYDALYQMD 360
Db      301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAEQLYDALYQMD 360
Qy      361 KQSLLEVTVDSLDFPKALYSDAATGYSSSSSTYSIVDAVKTPADGFVSIIVETHAASNG 420
Db      361 KQSLLEVTVDSLDFPKALYSDAATGYSSSSSTYSIVDAVKTPADGFVSIIVETHAASNG 420
Qy      421 SMSEQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Db      421 SMSEQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Qy      481 TYSQVTVTSWPSIVATGTTTATPTGSGSVTSTSKTTATATASKTSTTRSGMSL 530
Db      481 TYSQVTVTSWPSIVATGTTTATPTGSGSVTSTSKTTATATASKTSTTRSGMSL 530
```

## RESULT 7

```
US-10-421-586-14
; Sequence 14, Application US/10421586
; Publication No. US20040002142A1
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne Roenfeldt
; APPLICANT: Svendsen, Allan
; APPLICANT: Pedersen, Henrik
; APPLICANT: Vind, Jesper
; APPLICANT: Hendriksen, Hanne Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OR INVENTION: Glucoamylase Variants
; FILE REFERENCE: 5967.210-US
; CURRENT APPLICATION NUMBER: US/10/421,586
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 14
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-10-421-586-14
```

```
Query Match      99.0%; Score 2742; DB 15; Length 640;
Best Local Similarity 99.6%; Pred. No. 1.4e-234;
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Db      1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Qy      61 VVASPSNDNDPYFYTWTRDSGLVLTLYDLFRNGDTSLSLTENYISAQAIVQGISNPSG 120
Db      61 VVASPSNDNDPYFYTWTRDSGLVLTLYDLFRNGDTSLSLTENYISAQAIVQGISNPSG 120
Qy      121 DLSSGAGLGEPPKFNVDETAYTGSWGRPQDGPALRATAMIGFGQWLNDNGYTSRATDIW 180
Db      121 DLSSGAGLGEPPKFNVDETAYTGSWGRPQDGPALRATAMIGFGQWLNDNGYTSRATDIW 180
Qy      181 PLVNDLSYYAQYWNQGYDLMEBVNGSSPFTIAVQHRALVEGSAFATAVGSSCWCDSQ 240
Db      181 PLVNDLSYYAQYWNQGYDLMEBVNGSSPFTIAVQHRALVEGSAFATAVGSSCWCDSQ 240
Qy      241 APELICYLQSFMTGSLFIANFDSRSRSGKDNATLLGSIHTFDPBAACDSTFQPCSPRALA 300
Db      241 APELICYLQSFMTGSLFIANFDSRSRSGKDNATLLGSIHTFDPBAACDSTFQPCSPRALA 300
Qy      301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAEQLYDALYQMD 360
```

Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360  
Qy 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIYDAVKTFPADGFSIYETHAASNG 420  
Db 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIYDAVKTFPADGFSIYETHAASNG 420  
Qy 421 SMSQYDSDGEBQLSARDLTWMSYALLTPANNRRNSVVPASMGETSASVPGCAATSAIG 480  
Db 421 SMSQYDSDGEBQLSARDLTWMSYALLTPANNRRNSVVPASMGETSASVPGCAATSAIG 480  
Qy 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTTR 530  
Db 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530

## RESULT 8

US-10-418-836-9  
; Sequence 9, Application US/10418836  
; Publication No. US20040018573A1  
; GENERAL INFORMATION:  
; APPLICANT: Power, Scott D.  
; APPLICANT: Wang, Hoaming  
; APPLICANT: Ward, Michael  
; TITLE OF INVENTION: Production of Functional Antibodies in  
; FILE REFERENCE: GC741-2  
; CURRENT APPLICATION NUMBER: US/10/418,836  
; CURRENT FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: US 60/373,889  
; PRIOR FILING DATE: 2002-04-18  
; PRIOR APPLICATION NUMBER: US 60/411,540  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: US 60/452,134  
; PRIOR FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: US 60/411,537  
; PRIOR FILING DATE: 2002-09-18  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 9  
; LENGTH: 743  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: fusion protein  
US-10-418-836-9

Query Match 98.7%; Score 2735; DB 15; Length 743;  
Best Local Similarity 99.6%; Pred. No. 7.2e-234;  
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60  
Db 1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60  
Qy 61 VVASSTDNPDYFYWTTRDSGLVLTLYDLPRNGDTSLSLTIENYISAQAIYQGISNPSG 120  
Db 61 VVASSTDNPDYFYWTTRDSGLVLTLYDLPRNGDTSLSLTIENYISAQAIYQGISNPSG 120  
Qy 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALRATAMIGFGQWLLDNGYSTATDIYW 180  
Db 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALRATAMIGFGQWLLDNGYSTATDIYW 180  
Qy 181 PLVRNDLSYVAQYNNQGTGYDLMEBNVSGSFFITIAVQHRALVEGSAFATAVSSCSWCDSQ 240  
Db 181 PLVRNDLSYVAQYNNQGTGYDLMEBNVSGSFFITIAVQHRALVEGSAFATAVSSCSWCDSQ 240  
Qy 241 APEILCYQSFWTGSFIIANFPDSSRGKANTLLGSIHFPDEAACDDSTFOPGSPRALA 300  
Db 241 APEILCYQSFWTGSFIIANFPDSSRGKANTLLGSIHFPDEAACDDSTFOPGSPRALA 300  
Qy 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360  
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360

Qy 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIYDAVKTFPADGFSIYETHAASNG 420  
Db 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIYDAVKTFPADGFSIYETHAASNG 420  
Qy 421 SMSQYDSDGEBQLSARDLTWMSYALLTPANNRRNSVVPASMGETSASVPGCAATSAIG 480  
Db 421 SMSQYDSDGEBQLSARDLTWMSYALLTPANNRRNSVVPASMGETSASVPGCAATSAIG 480  
Qy 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTTR 529  
Db 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTYKR 529

## RESULT 9

US-10-418-836-13  
; Sequence 13, Application US/10418836  
; Publication No. US20040018573A1  
; GENERAL INFORMATION:  
; APPLICANT: Power, Scott D.  
; APPLICANT: Wang, Hoaming  
; APPLICANT: Ward, Michael  
; TITLE OF INVENTION: Production of Functional Antibodies in  
; FILE REFERENCE: GC741-2  
; CURRENT APPLICATION NUMBER: US/10/418,836  
; CURRENT FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: US 60/373,889  
; PRIOR FILING DATE: 2002-04-18  
; PRIOR APPLICATION NUMBER: US 60/411,540  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: US 60/452,134  
; PRIOR FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: US 60/411,537  
; PRIOR FILING DATE: 2002-09-18  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 13  
; LENGTH: 763  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: fusion protein  
US-10-418-836-13

Query Match 98.7%; Score 2735; DB 15; Length 763;  
Best Local Similarity 99.6%; Pred. No. 7.5e-234;  
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60  
Db 1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60  
Qy 61 VVASSTDNPDYFYWTTRDSGLVLTLYDLPRNGDTSLSLTIENYISAQAIYQGISNPSG 120  
Db 61 VVASSTDNPDYFYWTTRDSGLVLTLYDLPRNGDTSLSLTIENYISAQAIYQGISNPSG 120  
Qy 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALRATAMIGFGQWLLDNGYSTATDIYW 180  
Db 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALRATAMIGFGQWLLDNGYSTATDIYW 180  
Qy 181 PLVRNDLSYVAQYNNQGTGYDLMEBNVSGSFFITIAVQHRALVEGSAFATAVSSCSWCDSQ 240  
Db 181 PLVRNDLSYVAQYNNQGTGYDLMEBNVSGSFFITIAVQHRALVEGSAFATAVSSCSWCDSQ 240  
Qy 241 APEILCYQSFWTGSFIIANFPDSSRGKANTLLGSIHFPDEAACDDSTFOPGSPRALA 300  
Db 241 APEILCYQSFWTGSFIIANFPDSSRGKANTLLGSIHFPDEAACDDSTFOPGSPRALA 300  
Qy 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360  
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360

QY 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420  
DB 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420  
QY 421 SMSQYDKSDGEOLASADLTWMSYAALLTANNRRNSVVPASMGETSASVGTCAATSAIG 480  
DB 421 SMSQYDKSDGEOLASADLTWMSYAALLTANNRRNSVVPASMGETSASVGTCAATSAIG 480  
QY 481 TYSSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATATASKSTYKR 529  
DB 481 TYSSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATATASKSTYKR 529

RESULT 10  
US-10-418-836-10  
; Sequence 10, Application US/10418836  
; Publication No. US20040018573A1  
; GENERAL INFORMATION:  
; APPLICANT: Power, Scott D.  
; APPLICANT: Wang, Huaming  
; APPLICANT: Ward, Michael  
; TITLE OF INVENTION: Production of Functional Antibodies in  
; TITLE OF INVENTION: Filamentous Fungi  
; FILE REFERENCE: GC741-2  
; CURRENT APPLICATION NUMBER: US/10/418, 836  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: US 60/373,889  
; PRIOR FILING DATE: 2002-04-18  
; PRIOR APPLICATION NUMBER: US 60/411,540  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: US 60/452,134  
; PRIOR FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: US 60/411,537  
; PRIOR FILING DATE: 2002-09-18  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 979  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: fusion protein  
US-10-418-836-10

Query Match 98.7%; Score 2735; DB 15; Length 979;  
Best Local Similarity 99.6%; Pred. No. 1,1e-233;  
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSFRSLALSGLVCTGLANVISKRATLDSWLSNEATYARFALINNIIGADGAWSGADSGI 60  
DB 1 MSFRSLALSGLVCTGLANVISKRATLDSWLSNEATYARFALINNIIGADGAWSGADSGI 60  
QY 61 VVASPSTNDPDPYFTYTWTRDSGLVLTIVDLFRNGDTSLSLTENYISAQAIIVQGISNPSG 120  
DB 61 VVASPSTNDPDPYFTYTWTRDSGLVLTIVDLFRNGDTSLSLTENYISAQAIIVQGISNPSG 120  
QY 121 DLSGAGLGEPRKFNVDETAYATGSGRQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180  
DB 121 DLSGAGLGEPRKFNVDETAYATGSGRQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180  
QY 181 PLYVNDLSIYAQYVNOGTGYDLMEEVNGSSFFTLAVQHRALVEGSAFATAVGSSCSWCDSQ 240  
DB 181 PLYVNDLSIYAQYVNOGTGYDLMEEVNGSSFFTLAVQHRALVEGSAFATAVGSSCSWCDSQ 240  
QY 241 APELICYLQSFWMGSSFLIANPDSRSGKDANTLLGSIHTFDPBAACDSTFQPCSPRALA 300  
DB 241 APELICYLQSFWMGSSFLIANPDSRSGKDANTLLGSIHTFDPBAACDSTFQPCSPRALA 300  
QY 301 NHKEVVDSPRSIYTLNDGLSDSEBAVAAGRYPEDITYNGNPMFLCTLLAAEQLYDALYQMD 360  
DB 301 NHKEVVDSPRSIYTLNDGLSDSEBAVAAGRYPEDITYNGNPMFLCTLLAAEQLYDALYQMD 360  
QY 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420

DB 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420  
QY 421 SMSQYDKSDGEOLASADLTWMSYAALLTANNRRNSVVPASMGETSASVGTCAATSAIG 480  
DB 421 SMSQYDKSDGEOLASADLTWMSYAALLTANNRRNSVVPASMGETSASVGTCAATSAIG 480  
QY 481 TYSSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATATASKSTYKR 529  
DB 481 TYSSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATATASKSTYKR 529

RESULT 11  
US-10-418-836-16  
; Sequence 16, Application US/10418836  
; Publication No. US20040018573A1  
; GENERAL INFORMATION:  
; APPLICANT: Power, Scott D.  
; APPLICANT: Wang, Huaming  
; APPLICANT: Ward, Michael  
; TITLE OF INVENTION: Production of Functional Antibodies in  
; TITLE OF INVENTION: Filamentous Fungi  
; FILE REFERENCE: GC741-2  
; CURRENT APPLICATION NUMBER: US/10/418, 836  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: US 60/373,889  
; PRIOR FILING DATE: 2002-04-18  
; PRIOR APPLICATION NUMBER: US 60/411,540  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: US 60/452,134  
; PRIOR FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: US 60/411,537  
; PRIOR FILING DATE: 2002-09-18  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 979  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: fusion protein  
US-10-418-836-16

Query Match 98.7%; Score 2735; DB 15; Length 979;  
Best Local Similarity 99.6%; Pred. No. 1,1e-233;  
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSFRSLALSGLVCTGLANVISKRATLDSWLSNEATYARFALINNIIGADGAWSGADSGI 60  
DB 1 MSFRSLALSGLVCTGLANVISKRATLDSWLSNEATYARFALINNIIGADGAWSGADSGI 60  
QY 61 VVASPSTNDPDPYFTYTWTRDSGLVLTIVDLFRNGDTSLSLTENYISAQAIIVQGISNPSG 120  
DB 61 VVASPSTNDPDPYFTYTWTRDSGLVLTIVDLFRNGDTSLSLTENYISAQAIIVQGISNPSG 120  
QY 121 DLSGAGLGEPRKFNVDETAYATGSGRQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180  
DB 121 DLSGAGLGEPRKFNVDETAYATGSGRQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180  
QY 181 PLYVNDLSIYAQYVNOGTGYDLMEEVNGSSFFTLAVQHRALVEGSAFATAVGSSCSWCDSQ 240  
DB 181 PLYVNDLSIYAQYVNOGTGYDLMEEVNGSSFFTLAVQHRALVEGSAFATAVGSSCSWCDSQ 240  
QY 241 APELICYLQSFWMGSSFLIANPDSRSGKDANTLLGSIHTFDPBAACDSTFQPCSPRALA 300  
DB 241 APELICYLQSFWMGSSFLIANPDSRSGKDANTLLGSIHTFDPBAACDSTFQPCSPRALA 300  
QY 301 NHKEVVDSPRSIYTLNDGLSDSEBAVAAGRYPEDITYNGNPMFLCTLLAAEQLYDALYQMD 360  
DB 301 NHKEVVDSPRSIYTLNDGLSDSEBAVAAGRYPEDITYNGNPMFLCTLLAAEQLYDALYQMD 360  
QY 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420

Db 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420  
QY 421 SMSSEYDSDGSDGRLSARDLTWSYALTLNRRNSVVPASWETSASSVPGCAATSAIG 480  
Db 421 SMSSEYDSDGSDGRLSARDLTWSYALTLNRRNSVVPASWETSASSVPGCAATSAIG 480  
QY 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSKTTATASKTSITTR 529  
Db 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSKTTATASKTSITTR 529

RESULT 12  
US-10-418-836-19  
; Sequence 19, Application US/10418836  
; Publication No. US20040018573A1  
; GENERAL INFORMATION:  
; APPLICANT: Power, Scott D.  
; APPLICANT: Wang, Huaming  
; APPLICANT: Ward, Michael  
; TITLE OF INVENTION: Production of Functional Antibodies in  
; FILE REFERENCE: GC741-2  
; CURRENT APPLICATION NUMBER: US/10/418,836  
; CURRENT FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: US 60/373,889  
; PRIOR FILING DATE: 2002-04-18  
; PRIOR APPLICATION NUMBER: US 60/411,540  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: US 60/452,134  
; PRIOR FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: US 60/411,537  
; PRIOR FILING DATE: 2002-09-18  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 738  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: fusion protein  
US-10-418-836-19

Query Match 98.2%; Score 2720.5; DB 15; Length 728;  
Best Local Similarity 98.5%; Pred. No. 1.4e-232;  
Matches 527; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 MSFRLSLALSGIVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60  
Db 1 MSFRLSLALSGIVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60  
QY 61 VVASSTNPNPYFYTWTRDSGLVLTLYDLFRNGDTSLSTIENTISAOAIYQGISNPSG 120  
Db 61 VVASSTNPNPYFYTWTRDSGLVLTLYDLFRNGDTSLSTIENTISAOAIYQGISNPSG 120  
QY 121 DLSGAGLGEKPFNVDETAAYTSGMGRPORDBPALRATAMIGFQWLLNGYTSATDLYW 180  
Db 121 DLSGAGLGEKPFNVDETAAYTSGMGRPORDBPALRATAMIGFQWLLNGYTSATDLYW 180  
QY 181 PLVRNDLSYVAQYNNQGYDLMEEVNGSSFPIAVQHRALVEGSAFATAVSSCSWCDSQ 240  
Db 181 PLVRNDLSYVAQYNNQGYDLMEEVNGSSFPIAVQHRALVEGSAFATAVSSCSWCDSQ 240  
QY 241 APELICYSQFWTGSFILANPDSRSKCDANTLLGSIHTFDEPAACDSTFQPCSPRALA 300  
Db 241 APELICYSQFWTGSFILANPDSRSKCDANTLLGSIHTFDEPAACDSTFQPCSPRALA 300  
QY 301 NHKEVVDSPRSIYTLNDGLSDSEAVAAGRYPEDTYNNGNPMFLCTLAABOLYDALYQMD 360  
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVAAGRYPEDTYNNGNPMFLCTLAABOLYDALYQMD 360  
QY 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420  
Db 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420

QY 421 SMSSEYDSDGSDGRLSARDLTWSYALTLNRRNSVVPASWETSASSVPGCAATSAIG 480  
Db 421 SMSSEYDSDGSDGRLSARDLTWSYALTLNRRNSVVPASWETSASSVPGCAATSAIG 480  
QY 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSKTTATASKTSITTR 534  
Db 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSKTTATASKTSITTR 534

RESULT 13  
US-10-418-836-30  
; Sequence 30, Application US/10418836  
; Publication No. US20040018573A1  
; GENERAL INFORMATION:  
; APPLICANT: Power, Scott D.  
; APPLICANT: Wang, Huaming  
; APPLICANT: Ward, Michael  
; TITLE OF INVENTION: Production of Functional Antibodies in  
; FILE REFERENCE: GC741-2  
; CURRENT APPLICATION NUMBER: US/10/418,836  
; CURRENT FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: US 60/373,889  
; PRIOR FILING DATE: 2002-04-18  
; PRIOR APPLICATION NUMBER: US 60/411,540  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: US 60/452,134  
; PRIOR FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: US 60/411,537  
; PRIOR FILING DATE: 2002-09-18  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 741  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: fusion protein  
US-10-418-836-30

Query Match 98.0%; Score 2716; DB 15; Length 741;  
Best Local Similarity 98.0%; Pred. No. 3.5e-232;  
Matches 527; Conservative 1; Mismatches 6; Indels 4; Gaps 1;

QY 1 MSFRLSLALSGIVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60  
Db 1 MSFRLSLALSGIVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60  
QY 61 VVASSTNPNPYFYTWTRDSGLVLTLYDLFRNGDTSLSTIENTISAOAIYQGISNPSG 120  
Db 61 VVASSTNPNPYFYTWTRDSGLVLTLYDLFRNGDTSLSTIENTISAOAIYQGISNPSG 120  
QY 121 DLSGAGLGEKPFNVDETAAYTSGMGRPORDBPALRATAMIGFQWLLNGYTSATDLYW 180  
Db 121 DLSGAGLGEKPFNVDETAAYTSGMGRPORDBPALRATAMIGFQWLLNGYTSATDLYW 180  
QY 181 PLVRNDLSYVAQYNNQGYDLMEEVNGSSFPIAVQHRALVEGSAFATAVSSCSWCDSQ 240  
Db 181 PLVRNDLSYVAQYNNQGYDLMEEVNGSSFPIAVQHRALVEGSAFATAVSSCSWCDSQ 240  
QY 241 APELICYSQFWTGSFILANPDSRSKCDANTLLGSIHTFDEPAACDSTFQPCSPRALA 300  
Db 241 APELICYSQFWTGSFILANPDSRSKCDANTLLGSIHTFDEPAACDSTFQPCSPRALA 300  
QY 301 NHKEVVDSPRSIYTLNDGLSDSEAVAAGRYPEDTYNNGNPMFLCTLAABOLYDALYQMD 360  
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVAAGRYPEDTYNNGNPMFLCTLAABOLYDALYQMD 360  
QY 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420  
Db 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420

Qy	Db	Qy	Db
422	421	481	481
SMEQYDKSPGEOELSRDLTWSYAALLTNRRNRNVVPSANGESASSVPETCAATGAIG	SMEQYDKSDGELSRDLTWSYAALLTNRRNRNVVPSANGESASSVPETCAATGAIG	TSYSVVTSPSPSVAAVGCTTTATPFGSGSVTSKTTATASKTS-----TTTBGML	TSYSVVTSPSPSVAAVGCTTTATPFGSGSVTSKTTATASKTDIQTOSTPSLL
480	480	534	538

RESULT 14  
US-10-418-836-38  
; Sequence 38, Application US/10418836  
; Publication No. US20040018573A1

```

1 GENERAL INFORMATION:
2 APPLICANT: Power, Scott D.
3 APPLICANT: Wang, Huang
4 APPLICANT: Ward, Michael
5 TITLE OF INVENTION: Production of Functional
6 TITLE OF INVENTION: Filamentous Fungi
7 FILE REFERENCE: GC/41-2
8 CURRENT APPLICATION NUMBER: US/10/418,836
9 CURRENT FILING DATE: 2003-04-17
10 PRIOR APPLICATION NUMBER: US 60/373,889
11 PRIOR FILING DATE: 2002-04-18
12 PRIOR APPLICATION NUMBER: US 60/411,540
13 PRIOR FILING DATE: 2002-09-18
14 PRIOR APPLICATION NUMBER: US 60/452,134
15 PRIOR FILING DATE: 2003-03-04
16 PRIOR APPLICATION NUMBER: US 60/411,537
17 PRIOR FILING DATE: 2002-09-18
18 NUMBER OF SEQ ID NOS: 40
19 SOFTWARE: FastSeq for Windows Version 4.0.0
20 SEQ ID NO 38
21 LENGTH: 972
22 TYPE: PRNT
23 ORGANISM: Artificial Sequence
24 FEATURE:
25 OTHER INFORMATION: fusion protein
26 US-10-418-836-38

```

Query Match	98.0%;	Score 2715.5;	DB 15;	Length 972;
Best Local Similarly	95.5%;	Pred. No. 5.8e-232;		
Matches 530;	Conservative 0;	Mismatches 4;	Indels 21;	Gaps 2;

Qy	1	MSFSLSLASGLVCTGLANVTSKATLDSMTLSNEVTARTALINNI	GADGAMV	SADSGDI	60
Db	1	MSFSLSLASGLVCTGLANVTSKATLDSMTLSNEVTARTALINNI	GADGAMV	SADSGDI	60
Qy	61	VVASPSTNDPYPYTTWRDGLVLTAVLDFRNGDTSLLSTIENTYISAQALVQGISNPSG			120
Db	61	VVASPSTNDPYPYTTWRDGLVLTAVLDFRNGDTSLLSTIENTYISAQALVQGISNPSG			120
Qy	121	DLSSGAGIGEEKFVNDETAATYGSWCRPORDBPALATAMIGFGOMLDNGYTRATDVIW			180
Db	121	DLSSGAGIGEEKFVNDETAATYGSWCRPORDBPALATAMIGFGOMLDNGYTRATDVIW			180
Qy	181	PLVANDLSYVAQYNNQGTGYDLMEENVGSSPFTIAVQHALLVEGSAFATAVSSCSWCDSQ			240
Db	181	PLVANDLSYVAQYNNQGTGYDLMEENVGSSPFTIAVQHALLVEGSAFATAVSSCSWCDSQ			240
Qy	241	APELLCTYQSFWTGSPFIILANFDSRSGKDANTLGSHTFPBEAACDDSTPOCSPRALA			300
Db	241	APELLCTYQSFWTGSPFIILANFDSRSGKDANTLGSHTFPBEAACDDSTPOCSPRALA			300
Qy	301	NHKEVNSFRSIYTLNDGLDSEAVAVGRYEPDTPYNNCPMPELCTILAAEQLYALXYOMD			360
Db	301	NHKEVNSFRSIYTLNDGLDSEAAVAVRYPEDTPYNNCPMPELCTILAAEQLYALXYOMD			360
Qy	361	KQGSLEVDVSLDFPKALYSDAATCTGYSSSSSTYSIYDAYKTFADGVSIVETHPAASNG			420
Db	361	KQGSLEVDVSLDFPKALYSDAATCTGYSSSSSTYSIYDAYKTFADGVSIVETHPAASNG			420
Qy	421	SMSBOYDKSDSEOLASRDLTWSYALALTANNRRNSVPAWSGETSASSGPGTCATSAIG			480

Db 421 SMEQIDKSDGGQLSARDLTWSYALLLRNNRRSSVYPASNGEISASASVGTCAISAIG 480  
 QY 481 TYSSTVTISWPSIVATGTTTTATPTGSGSVTSISKTTAAS-----KT 524  
 Db 481 TYSSTVTISWPSIVATGTTTTATPTGSGSVTSISKTTAATASIKROVOLAQSGPLVKP 540  
 QY 525 S-----TTRSGMSL 534  
 Db 541 SETLSTLCTVSGPRL 555

RESULT 15  
US-10-418-836-26  
; Sequence 26, Application US/10418836  
; Publication No. US20040018573A1

```

1 GENERAL INFORMATION:
2 APPLICANT: Power, Scott D.
3 APPLICANT: Wang, Huang
4 APPLICANT: Ward, Michael
5 TITLE OF INVENTION: Production of Functional
6 FILE OF INVENTION: Filamentous Fungi
7 TITLE REFERENCE: GC741-2
8 CURRENT APPLICATION NUMBER: US/10/418,836
9 CURRENT FILING DATE: 2003-04-17
10 PRIOR APPLICATION NUMBER: US 60/373,089
11 PRIOR FILING DATE: 2002-04-18
12 PRIOR APPLICATION NUMBER: US 60/411,540
13 PRIOR FILING DATE: 2002-09-18
14 PRIOR APPLICATION NUMBER: US 60/452,134
15 PRIOR FILING DATE: 2003-03-04
16 PRIOR APPLICATION NUMBER: US 60/411,537
17 PRIOR FILING DATE: 2002-09-18
18 NUMBER OF SEQ ID NOS: 40
19 SOFTWARE: FastSeq For Windows Version 4.0.
20 SEQ ID NO 26
21 LENGTH: 739
22 TYPE: PRT
23 ORGANISM: Artificial Sequence
24 FEATURE:
25 OTHER INFORMATION: fusion protein
26 US-10-418-836-26

```

Query Match	98.0%	Score 2715;	DB 15	Length 739;
Best Local Similarity	99.8%	Pred. No. 4.3e-232;		
Matches 522;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MSFRSLALSGVCTGLANVISKRATLTDMSLNEATVARTAILNINIGADGAWVSGADSGI	60
Dh	1	MSFRSLALSGVCTGLANVISKRATLTDMSLNEATVARTAILNINIGADGAWVSGADSGI	60
Qy	61	VVASPSTNDPDYFTYWTTRDSGLVLTALVDLFRNGDTSILSTIENYISAQAIIVOGISNPSG	120
Dh	61	VVASPSTNDPDYFTYWTTRDSGLVLTALVDLFRNGDTSILSTIENYISAQAIIVOGISNPSG	120
Qy	121	DLSSGGLGEPEKPNVDETYYTGSMSRPRDGPALRATMTMGQWLLDNGYSTANDIYW	180
Dh	121	DLSSGGLGEPEKPNVDETYYTGSMSRPRDGPALRATMTMGQWLLDNGYSTANDIYW	180
Qy	181	PLVRNRLNSTVAAQNMOTGYDLMEEVNGSSPFTIANQHRALVEGSAFPAVAGSSCSMCDSQ	240
Dh	181	PLVRNRLNSTVAAQNMOTGYDLMEEVNGSSPFTIANQHRALVEGSAFPAVAGSSCSMCDSQ	240
Qy	241	APRILCYLOSFWTGSFTILANPESSRSRGDANTLLGSIHTPEDEAACDSTFPOCSPRALA	300
Dh	241	APRILCYLOSFWTGSFTILANPESSRSRGDANTLLGSIHTPEDEAACDSTFPOCSPRALA	300
Qy	301	NHKEVVDSEFRSIYTLINDGISDSEAAVAVGRYPEDTYNGNPMWFLCTLLAAEQLDALYQMD	360
Dh	301	NHKEVVDSEFRSIYTLINDGISDSEAAVAVGRYPEDTYNGNPMWFLCTLLAAEQLDALYQMD	360
Qy	361	KQGSLEVDVSJLDFRKALYSDAATGTYSSSSSTYSIYDAVAKTYPADGFSIVETHAASNG	420

Db	361	KQGSLEVTDVSLDPFKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIETHAASNG	420
Oy	421	SMSEQYDKSDGBOQSARDLTWSYALLTANNRRNSVVPASWGETGASSVPGTCAATSAG	480
Db	421	SMSEQYDKSDGBOQSARDLTWSYALLTANNRRNSVVPASWGETSASSVPGTCAATSAG	480
Oy	481	TYSSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATASK	523
Db	481	TYSSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATASQ	523

Search completed: June 17, 2004, 17:53:29  
 Job time : 51 secs

2110 PAGE LEFT BLANK



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 17, 2004, 17:44:49 ; Search time 23 Seconds  
(without alignments)  
1198.620 Million cell updates/sec

Title: US-10-038-723-2  
Perfect score: 2771  
Sequence: 1 MSFSLALSGLVCTGLANV.....SKTTATASKSTTTRSGMSL 534

Scoring table: BLASTSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/6C\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2771	100.0	534	US-09-198-672-2	Sequence 2, Appli
2	2771	100.0	534	US-09-199-290-9	Sequence 9, Appli
3	2771	100.0	534	US-09-632-392-2	Sequence 2, Appli
4	2771	100.0	534	US-09-455-679-1	Sequence 1, Appli
5	2771	100.0	534	US-09-351-814-2	Sequence 2, Appli
6	2771	100.0	534	US-09-821-616-9	Sequence 9, Appli
7	2742	99.0	640	US-09-351-814-13	Sequence 13, Appli
8	2628	94.8	616	US-09-236-063-1	Sequence 1, Appli
9	1689	61.0	618	US-09-199-290-34	Sequence 34, Appli
10	1689	61.0	618	US-09-821-616-34	Sequence 34, Appli
11	1674.5	60.4	591	US-09-199-290-7	Sequence 7, Appli
12	1674.5	60.4	591	US-09-821-616-7	Sequence 7, Appli
13	1615.5	58.3	630	US-09-704-449-2	Sequence 2, Appli
14	1577	56.9	626	US-08-596-300A-7	Sequence 7, Appli
15	1577	56.9	626	US-08-596-300A-14	Sequence 14, Appli
16	1425.5	51.4	581	US-09-534-407-3	Sequence 3, Appli
17	1425.5	51.4	581	US-09-999-201B-4	Sequence 4, Appli
18	1321	47.7	616	US-08-385-370-2	Sequence 2, Appli
19	1321	47.7	616	US-08-385-370-4	Sequence 4, Appli
20	589	21.3	806	US-08-270-076A-11	Sequence 11, Appli
21	160	5.8	35	US-09-199-290-4	Sequence 4, Appli
22	160	5.8	35	US-09-821-616-4	Sequence 4, Appli
23	156.5	5.6	2137	US-09-134-001C-4463	Sequence 4463, Ap
24	142	5.1	536	US-09-463-712C-10	Sequence 10, Appli
25	136	4.9	1026	US-08-614-377A-7	Sequence 7, Appli
26	136	4.9	1026	US-09-142-648B-7	Sequence 7, Appli
27	133	4.8	1529	US-08-728-470-10	Sequence 10, Appli

28	133	4.8	1529	3	US-08-719-641-10	Sequence 10, Appli
29	132	4.8	1426	3	US-09-136-574A-43	Sequence 43, Appli
30	130.5	4.7	1537	1	US-08-328-267A-2	Sequence 2, Appli
31	130	4.7	1600	2	US-08-617-687-10	Sequence 10, Appli
32	129	4.7	1026	1	US-08-194-290-7	Sequence 7, Appli
33	128.5	4.6	673	4	US-09-107-532A-5134	Sequence 5134, Ap
34	128	4.6	1721	3	US-08-700-651-5	Sequence 5, Appli
35	128	4.6	1721	3	US-08-928-361B-6	Sequence 6, Appli
36	128	4.6	1721	4	US-09-588-995A-6	Sequence 6, Appli
37	127.5	4.6	751	6	5457037-5	Patent No. 5457037
38	127.5	4.6	776	3	US-09-346-237-7	Sequence 7, Appli
39	127	4.6	630	2	US-08-797-366-3	Sequence 3, Appli
40	127	4.6	630	2	US-08-956-268-3	Sequence 3, Appli
41	127	4.6	906	4	US-09-717-364A-15	Sequence 15, Appli
42	126.5	4.6	776	3	US-09-346-237-4	Sequence 4, Appli
43	126	4.5	288	4	US-09-216-393B-341	Sequence 341, App
44	126	4.5	288	4	US-09-216-393B-344	Sequence 344, App
45	124.5	4.5	894	3	US-08-362-525-22	Sequence 22, Appli

ALIGNMENTS

```
RESULT 1
US-09-198-672-2
; Sequence 2, Application US/09198672
; Patent No. 6129788
; GENERAL INFORMATION:
; APPLICANT: Liaw, Gin
; APPLICANT: Pedersen, Sven
; TITLE OF INVENTION: A Method of Producing Saccharide
; TITLE OF INVENTION: Preparations
; FILE REFERENCE: 5318,200-US
; CURRENT APPLICATION NUMBER: US/09/198, 672
; CURRENT FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus Niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-198-672-2
Query Match 100.0%; Score 2771; DB 3; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.1e-231;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFSLALSGLVCTGLANVISKRAATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60
DB 1 MSFSLALSGLVCTGLANVISKRAATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60
QY 61 VVAPSPDNPFFYFTWTRDGLVTKTLVDLFRNDTSLSTIEYISQAIVQGISNSG 120
DB 61 VVAPSPDNPFFYFTWTRDGLVTKTLVDLFRNDTSLSTIEYISQAIVQGISNSG 120
QY 121 DLSSGAGGEPKFNVDATAVYSGNGRPQDOPALRATAMIFGQWLDLNGYSTATDVIW 180
DB 121 DLSSGAGGEPKFNVDATAVYSGNGRPQDOPALRATAMIFGQWLDLNGYSTATDVIW 180
QY 181 PLYVNDLSYVAQYNQGYDIMEEVNGSSFFTIAVQIRALVSGSAFATVAGSSCSWCDSQ 240
DB 181 PLYVNDLSYVAQYNQGYDIMEEVNGSSFFTIAVQIRALVSGSAFATVAGSSCSWCDSQ 240
QY 241 APELICVQSGFWTGSFLLANPDSRSRGKANDTLTIGSIHTTPREACDSTFQPCSPRALA 300
DB 241 APELICVQSGFWTGSFLLANPDSRSRGKANDTLTIGSIHTTPREACDSTFQPCSPRALA 300
QY 301 NHRKVVNSFRSITYTLNDGLDSEBAVAVGRVPEDTTYNGNPFILCTLLAAEQLYALYQWD 360
DB 301 NHRKVVNSFRSITYTLNDGLDSEBAVAVGRVPEDTTYNGNPFILCTLLAAEQLYALYQWD 360
```

```
Db 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTLAAEQLYDALYQMD 360
Qy 361 KQGSLEVTDVSLDFKALYSDAATGTSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Db 361 KQGSLEVTDVSLDFKALYSDAATGTSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Qy 421 SMSBOYKSDGEQUSARDLWTSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Db 421 SMSBOYKSDGEQUSARDLWTSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Qy 481 TYSVSVTSMPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
Db 481 TYSVSVTSMPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

```
RESULT 2
US-09-199-290-9
; Sequence 9, Application US/09199290
; Patent No. 6255084
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Rudy
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279, 200-US
; CURRENT APPLICATION NUMBER: US/09/199, 290
; EARLIER FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 1557/97
; EARLIER FILING DATE: 1997-12-30
; EARLIER APPLICATION NUMBER: 0925/98
; EARLIER FILING DATE: 1998-07-10
; EARLIER APPLICATION NUMBER: 60/070, 746
; EARLIER FILING DATE: 1998-01-08
; EARLIER APPLICATION NUMBER: 60/094, 344
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/979, 673
; EARLIER FILING DATE: 1997-11-26
; EARLIER APPLICATION NUMBER: 09/107, 657
; EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-199-290-9
```

```
Query Match 100.0%; Score 2771; DB 3; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.1e-231;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Db 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Qy 61 VVASPSTNDNDYFTYTWTRDSGLVYKTLVDLFRNGDTSLSTIENYISAQALVQGISNPSG 120
Db 61 VVASPSTNDNDYFTYTWTRDSGLVYKTLVDLFRNGDTSLSTIENYISAQALVQGISNPSG 120
Qy 121 DLSSGAGLGEPRKFNVDETATYTGSMGRPORDGPALRATAMIGFGOMLNDNGTSTATDIW 180
Db 121 DLSSGAGLGEPRKFNVDETATYTGSMGRPORDGPALRATAMIGFGOMLNDNGTSTATDIW 180
Qy 181 PLVNDLSYVAQYVNGQGYDLMEEVNGSSFFTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
Db 181 PLVNDLSYVAQYVNGQGYDLMEEVNGSSFFTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
Qy 241 APELICLOSFWTGSFIIANFDSRSRSGKDANTLLGSHTFDPEAACDSTFQPCSPRALA 300
Db 241 APELICLOSFWTGSFIIANFDSRSRSGKDANTLLGSHTFDPEAACDSTFQPCSPRALA 300
Qy 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTLAAEQLYDALYQMD 360
Db 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTLAAEQLYDALYQMD 360
```

```
Qy 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTLAAEQLYDALYQMD 360
Db 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTLAAEQLYDALYQMD 360
Qy 361 KQGSLEVTDVSLDFKALYSDAATGTSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Db 361 KQGSLEVTDVSLDFKALYSDAATGTSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Qy 421 SMSBOYKSDGEQUSARDLWTSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Db 421 SMSBOYKSDGEQUSARDLWTSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Qy 481 TYSVSVTSMPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
Db 481 TYSVSVTSMPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

```
RESULT 3
US-09-632-392-2
; Sequence 2, Application US/09632392
; Patent No. 6303346
; GENERAL INFORMATION:
; APPLICANT: Liaw, Gin
; APPLICANT: Pedersen, Sven
; APPLICANT: Hendriksen, Sven
; TITLE OF INVENTION: A Method of Producing Saccharide
; PREPARATIONS
; FILE REFERENCE: 5318, 200-US
; CURRENT APPLICATION NUMBER: US/09/632, 392
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 09/198, 672
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-632-392-2
```

```
Query Match 100.0%; Score 2771; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.1e-231;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Db 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Qy 61 VVASPSTNDNDYFTYTWTRDSGLVYKTLVDLFRNGDTSLSTIENYISAQALVQGISNPSG 120
Db 61 VVASPSTNDNDYFTYTWTRDSGLVYKTLVDLFRNGDTSLSTIENYISAQALVQGISNPSG 120
Qy 121 DLSSGAGLGEPRKFNVDETATYTGSMGRPORDGPALRATAMIGFGOMLNDNGTSTATDIW 180
Db 121 DLSSGAGLGEPRKFNVDETATYTGSMGRPORDGPALRATAMIGFGOMLNDNGTSTATDIW 180
Qy 181 PLVNDLSYVAQYVNGQGYDLMEEVNGSSFFTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
Db 181 PLVNDLSYVAQYVNGQGYDLMEEVNGSSFFTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
Qy 241 APELICLOSFWTGSFIIANFDSRSRSGKDANTLLGSHTFDPEAACDSTFQPCSPRALA 300
Db 241 APELICLOSFWTGSFIIANFDSRSRSGKDANTLLGSHTFDPEAACDSTFQPCSPRALA 300
Qy 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTLAAEQLYDALYQMD 360
Db 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTLAAEQLYDALYQMD 360
Qy 361 KQGSLEVTDVSLDFKALYSDAATGTSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
```

Db 361 KQGSLEVTVDLDFKALYSDAATGYSSSSSTYSIVDAVKTADFVSIYETHAASNG 420  
Qy 421 SMSSEYDSDGDEQSLARDLTWSYALLTANNRRNSVVPASWGETSASSVPGCAATSAIG 480  
Db 421 SMSSEYDSDGDEQSLARDLTWSYALLTANNRRNSVVPASWGETSASSVPGCAATSAIG 480  
Qy 481 TYSSTVTSWPSIYATGGTTTATPTGSGSVTSKTKTATASKSTTTTRSGMSL 534  
Db 481 TYSSTVTSWPSIYATGGTTTATPTGSGSVTSKTKTATASKSTTTTRSGMSL 534

RESULT 4  
US-09-455-679-1  
Sequence 1, Application US/09455679

Patent No. 6329186  
GENERAL INFORMATION:  
APPLICANT: Nielsen, Bjarne R.  
APPLICANT: Svendsen, Allan  
APPLICANT: Bojlsen, Kirsten  
APPLICANT: Vind, Jesper  
APPLICANT: Pedersen, Henrik  
TITLE OF INVENTION: Glucoamyliase With N-Terminal Extensions  
FILE REFERENCE: 5691.200-US  
CURRENT APPLICATION NUMBER: US/09/455,679  
EARLIER FILING DATE: 1999-12-07  
EARLIER APPLICATION NUMBER: PA 1998 01616  
EARLIER FILING DATE: 1998-12-07  
EARLIER APPLICATION NUMBER: PA 1999 00409  
EARLIER FILING DATE: 1999-03-24  
EARLIER APPLICATION NUMBER: 60/111,674  
EARLIER FILING DATE: 1998-12-10  
EARLIER APPLICATION NUMBER: 60/126,740  
EARLIER FILING DATE: 1999-03-29  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 534  
TYPE: PRT  
ORGANISM: Aspergillus niger  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: (1)...(24)  
US-09-455-679-1

Query Match 100.0%; Score 2771; DB 4; Length 534;  
Best Local Similarity 100.0%; Pred. No. 1,1e-231;  
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTALINNIGADGAMVSGADSGI 60  
Db 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTALINNIGADGAMVSGADSGI 60  
Qy 61 VVASBSTNDPYPFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTYISQAQIVQGISNPSG 120  
Db 61 VVASBSTNDPYPFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTYISQAQIVQGISNPSG 120  
Qy 121 DLSSGAGLGEPEKFNVDETAYTGSWGRPORDGPALRATAMIGFGQWLDNGYTSRTADIY 180  
Db 121 DLSSGAGLGEPEKFNVDETAYTGSWGRPORDGPALRATAMIGFGQWLDNGYTSRTADIY 180  
Qy 181 PLVNRDLSYVAQYNNQGYDLMEEVNGSSFPTIAVQHRALVEGSAFATAVGSSCSWCDSQ 240  
Db 181 PLVNRDLSYVAQYNNQGYDLMEEVNGSSFPTIAVQHRALVEGSAFATAVGSSCSWCDSQ 240  
Qy 241 APELICYSQSFWTGSFILANPDSRSRGDANTLLGSIHTFDEAACDSTFQPCSPRALA 300  
Db 241 APELICYSQSFWTGSFILANPDSRSRGDANTLLGSIHTFDEAACDSTFQPCSPRALA 300  
Qy 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABEQLYDALYQMD 360  
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABEQLYDALYQMD 360

Qy 361 KQGSLEVTVDLDFKALYSDAATGYSSSSSTYSIVDAVKTADFVSIYETHAASNG 420  
Db 361 KQGSLEVTVDLDFKALYSDAATGYSSSSSTYSIVDAVKTADFVSIYETHAASNG 420  
Qy 421 SMSSEYDSDGDEQSLARDLTWSYALLTANNRRNSVVPASWGETSASSVPGCAATSAIG 480  
Db 421 SMSSEYDSDGDEQSLARDLTWSYALLTANNRRNSVVPASWGETSASSVPGCAATSAIG 480  
Qy 481 TYSSTVTSWPSIYATGGTTTATPTGSGSVTSKTKTATASKSTTTTRSGMSL 534  
Db 481 TYSSTVTSWPSIYATGGTTTATPTGSGSVTSKTKTATASKSTTTTRSGMSL 534

RESULT 5  
US-09-351-814-2  
Sequence 2, Application US/09351814

Patent No. 6352851  
GENERAL INFORMATION:  
APPLICANT: Nielsen, Bjarne Roenfeldt  
APPLICANT: Svendsen, Allan  
APPLICANT: Pedersen, Henrik  
APPLICANT: Vind, Jesper  
APPLICANT: Hendriksen, Hanne Vang  
APPLICANT: Frandsen, Torben Peter  
TITLE OF INVENTION: Glucoamyliase Variants  
FILE REFERENCE: 5636.200-US  
CURRENT APPLICATION NUMBER: US/09/351,814  
EARLIER FILING DATE: 1999-07-12  
EARLIER APPLICATION NUMBER: PA 1998 00937  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: PA 1998 01667  
EARLIER FILING DATE: 1998-12-17  
EARLIER APPLICATION NUMBER: 60/093,528  
EARLIER FILING DATE: 1998-07-21  
EARLIER APPLICATION NUMBER: 60/115,545  
EARLIER FILING DATE: 1999-01-12  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 534  
TYPE: PRT  
ORGANISM: Aspergillus niger  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: (1)...(24)  
US-09-351-814-2

Query Match 100.0%; Score 2771; DB 4; Length 534;  
Best Local Similarity 100.0%; Pred. No. 1,1e-231;  
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTALINNIGADGAMVSGADSGI 60  
Db 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTALINNIGADGAMVSGADSGI 60  
Qy 61 VVASBSTNDPYPFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTYISQAQIVQGISNPSG 120  
Db 61 VVASBSTNDPYPFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTYISQAQIVQGISNPSG 120  
Qy 121 DLSSGAGLGEPEKFNVDETAYTGSWGRPORDGPALRATAMIGFGQWLDNGYTSRTADIY 180  
Db 121 DLSSGAGLGEPEKFNVDETAYTGSWGRPORDGPALRATAMIGFGQWLDNGYTSRTADIY 180  
Qy 181 PLVNRDLSYVAQYNNQGYDLMEEVNGSSFPTIAVQHRALVEGSAFATAVGSSCSWCDSQ 240  
Db 181 PLVNRDLSYVAQYNNQGYDLMEEVNGSSFPTIAVQHRALVEGSAFATAVGSSCSWCDSQ 240  
Qy 241 APELICYSQSFWTGSFILANPDSRSRGDANTLLGSIHTFDEAACDSTFQPCSPRALA 300  
Db 241 APELICYSQSFWTGSFILANPDSRSRGDANTLLGSIHTFDEAACDSTFQPCSPRALA 300  
Qy 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABEQLYDALYQMD 360  
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABEQLYDALYQMD 360

```
Db 301 NHKEVDSFRSIYTLNDGLSDSEAVAGRYPEPTYNGNPMWFLCTLAAEQYDALYQMD 360
Qy 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Db 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Qy 421 SMSBQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVPCTCAATSAIG 480
Db 421 SMSBQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVPCTCAATSAIG 480
Qy 481 TYSSTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
Db 481 TYSSTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

## RESULT 6

```
US-09-821-616-9
; Sequence 9, Application US/09821616
; Patent No. 6620924
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Ruby
; APPLICANT: Lembeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279.200-US
; CURRENT APPLICATION NUMBER: US/09/821.616
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/199,290
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1557/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 0925/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/070,746
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/094,344
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/979,673
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/107,657
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-821-616-9
```

```
Query Match 100.0%; Score 2771; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.1e-211;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFRLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60
Db 1 MSFRLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60
Qy 61 VVASPSTDNPDYFYTWTRDSGLVLTIVDLFRNGDTSLSLTIEHYISAQAIIVQGISNPSG 120
Db 61 VVASPSTDNPDYFYTWTRDSGLVLTIVDLFRNGDTSLSLTIEHYISAQAIIVQGISNPSG 120
Qy 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGQWLLDNGYTSATDIYW 180
Db 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGQWLLDNGYTSATDIYW 180
Qy 181 PLYVNDLSYVAQYWNQGYDLMEEVNGSSPFTIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
Db 181 PLYVNDLSYVAQYWNQGYDLMEEVNGSSPFTIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
Qy 241 APEILCYLQSFMTGSLFLANFDSRSRGKDNANTLIGSIHTFDPEAACDDSTFQPCSPRALA 300
```

```
Db 241 APEILCYLQSFMTGSLFLANFDSRSRGKDNANTLIGSIHTFDPEAACDDSTFQPCSPRALA 300
Qy 301 NHKEVDSFRSIYTLNDGLSDSEAVAGRYPEPTYNGNPMWFLCTLAAEQYDALYQMD 360
Db 301 NHKEVDSFRSIYTLNDGLSDSEAVAGRYPEPTYNGNPMWFLCTLAAEQYDALYQMD 360
Qy 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Db 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Qy 421 SMSBQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVPCTCAATSAIG 480
Db 421 SMSBQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVPCTCAATSAIG 480
Qy 481 TYSSTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
Db 481 TYSSTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

## RESULT 7

```
US-09-351-814-13
; Sequence 13, Application US/09351814
; Patent No. 6332851
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne Roenfeldt
; APPLICANT: Svendsen, Allan
; APPLICANT: Pedersen, Henrik
; APPLICANT: Wind, Jesper
; APPLICANT: Hendriksen, Hanne Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OF INVENTION: Glucoamylase Variants
; FILE REFERENCE: 5636.200-US
; CURRENT APPLICATION NUMBER: US/09/351,814
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: PA 1998 00937
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: PA 1998 01667
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/093,528
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/115,545
; EARLIER FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 640
; TYPE: PRT
; ORGANISM: ASPERGILLUS NIGER
US-09-351-814-13
```

```
Query Match 99.0%; Score 2742; DB 4; Length 640;
Best Local Similarity 99.6%; Pred. No. 4.9e-229;
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSFRLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60
Db 1 MSFRLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60
Qy 61 VVASPSTDNPDYFYTWTRDSGLVLTIVDLFRNGDTSLSLTIEHYISAQAIIVQGISNPSG 120
Db 61 VVASPSTDNPDYFYTWTRDSGLVLTIVDLFRNGDTSLSLTIEHYISAQAIIVQGISNPSG 120
Qy 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGQWLLDNGYTSATDIYW 180
Db 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGQWLLDNGYTSATDIYW 180
Qy 181 PLYVNDLSYVAQYWNQGYDLMEEVNGSSPFTIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
Db 181 PLYVNDLSYVAQYWNQGYDLMEEVNGSSPFTIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
Qy 241 APEILCYLQSFMTGSLFLANFDSRSRGKDNANTLIGSIHTFDPEAACDDSTFQPCSPRALA 300
```

Db 241 APEILCYLOSFWTGSFILANFDSRSRSGKANTLLGSIHTFDEAACDDSTFQPCSPRALA 300  
 QY 301 NHKEVVDSPRSIYTLNDGSDSEAVAVGRYPEDTYNGNPMFLCTLLAAAEQLYDALYQWD 360  
 Db 301 NHKEVVDSPRSIYTLNDGSDSEAVAVGRYPEDTYNGNPMFLCTLLAAAEQLYDALYQWD 360  
 QY 361 KQGSLEVTVDVSLDFPKALYSDAATGYSSSSSTYSIYDAVKTFPADGFSIYETHAASNG 420  
 Db 361 KQGSLEVTVDVSLDFPKALYSDAATGYSSSSSTYSIYDAVKTFPADGFSIYETHAASNG 420  
 QY 421 SMSSEYDSDGSEQLSARDLTWSYALLTANNRNSVVPASWGETSASVPGTCAATSAIG 480  
 Db 421 SMSSEYDSDGSEQLSARDLTWSYALLTANNRNSVVPASWGETSASVPGTCAATSAIG 480  
 QY 481 TYSSVTYVSWPSIVATGTTTATPTGSGSVTSTSKTTATKSTSTTS 530  
 Db 481 TYSSVTYVSWPSIVATGTTTATPTGSGSVTSTSKTTATKSTSTTS 530

## RESULT 8

US-09-236-063-1  
 ; Sequence 1, Application US/09236063  
 ; Patent No. 6537792

## GENERAL INFORMATION:

APPLICANT: Allen, Martin  
 APPLICANT: Fang, Tsuei-Yun  
 APPLICANT: Li, Yuxing  
 APPLICANT: Liu, Hsuan-Liang  
 APPLICANT: Chen, Haul-Mei  
 APPLICANT: Coutinho, Pedro  
 APPLICANT: Hanzatko, Richard  
 APPLICANT: Ford, Clark  
 TITLE OF INVENTION: PROTEIN ENGINEERING OF GLUCOAMYLASE TO  
 TITLE OF INVENTION: INCREASE PH OPTIMUM, SUBSTRATE SPECIFICITY AND  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Kohn & Associates  
 STREET: 30500 No. 6537792Western Hwy.  
 CITY: Farmington Hills  
 STATE: Michigan  
 COUNTRY: US  
 ZIP: 48334  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/236,063  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kohn, Kenneth I.  
 REGISTRATION NUMBER: 30,955  
 REFERENCE/DOCKET NUMBER: 0812.00001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (248) 539-5050  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 616 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Aspergillus  
 US-09-236-063-1

Query Match 94.8%; Score 2628; DB 4; Length 616;  
 Best Local Similarity 99.6%; Pred. No. 3,5e-219;

Matches 504; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 ATLDWLSNENAVTARTALLNIGADGAVSGADSGIYVASTNDPDTFYTTTRSGVL 84  
 Db 1 ATLDWLSNENAVTARTALLNIGADGAVSGADSGIYVASTNDPDTFYTTTRSGVL 60  
 QY 85 KTLVDFPNSGTSLSTIENYISAQAIYQGISNPSGDLSSGAGLGPFPNVDETAYTGSW 144  
 Db 61 KTLVDFPNSGTSLSTIENYISAQAIYQGISNPSGDLSSGAGLGPFPNVDETAYTGSW 120  
 QY 145 GRPQDGPALBATAWIGFGWLLDNGYSTATDIWPLVRDLASYAQQYMNQGYDLWE 204  
 Db 121 GRPQDGPALBATAWIGFGWLLDNGYSTATDIWPLVRDLASYAQQYMNQGYDLWE 180  
 QY 205 VNGSSFFITIAVQHRALVGSAPATVSGSSCWCSQAEIILCYLOSFWTGSFILANFDS 264  
 Db 181 VNGSSFFITIAVQHRALVGSAPATVSGSSCWCSQAEIILCYLOSFWTGSFILANFDS 240  
 QY 265 RSGKANTLLGSIHTFDEAACDDSTFQPCSPRALANKEVVDSPRSIYTLNDGSDSPA 324  
 Db 241 RSGKANTLLGSIHTFDEAACDDSTFQPCSPRALANKEVVDSPRSIYTLNDGSDSPA 300  
 QY 325 VAVGRYPEDTYNGNPMFLCTLLAAAEQLYDALYQWDKQGSLEVTVDVSLDFPKALYSDAAT 384  
 Db 301 VAVGRYPEDTYNGNPMFLCTLLAAAEQLYDALYQWDKQGSLEVTVDVSLDFPKALYSDAAT 360  
 QY 385 GTYSSSSSTYSIYDAVKTFPADGFSIYETHAASNGSSEQYDKSDGQLSARDLTWSYA 444  
 Db 361 GTYSSSSSTYSIYDAVKTFPADGFSIYETHAASNGSSEQYDKSDGQLSARDLTWSYA 420  
 QY 445 ALLTANNRNSVVPASWGETSASVPGTCAATSAIGTSSVTYVSWPSIVATGTTTAT 504  
 Db 421 ALLTANNRNSVVPASWGETSASVPGTCAATSAIGTSSVTYVSWPSIVATGTTTAT 480  
 QY 505 PTGSGSVTSTSKTTATKSTSTTS 530  
 Db 481 PTGSGSVTSTSKTTATKSTSTTS 506

## RESULT 9

US-09-199-290-34  
 ; Sequence 34, Application US/09199290  
 ; Patent No. 6255084

## GENERAL INFORMATION:

APPLICANT: Nielsen, Bjarne R.  
 APPLICANT: Nielsen, Ruby  
 APPLICANT: Lembeck, Jan  
 TITLE OF INVENTION: Thermostable Glucoamylase  
 FILE REFERENCE: 5279 200-US  
 CURRENT APPLICATION NUMBER: US/09/199,290  
 CURRENT FILING DATE: 1998-11-24  
 EARLIER APPLICATION NUMBER: 1557/97  
 EARLIER FILING DATE: 1997-12-30  
 EARLIER APPLICATION NUMBER: 0925/98  
 EARLIER FILING DATE: 1998-07-10  
 EARLIER APPLICATION NUMBER: 60/070,746  
 EARLIER FILING DATE: 1998-01-08  
 EARLIER APPLICATION NUMBER: 60/094,344  
 EARLIER FILING DATE: 1998-07-28  
 EARLIER APPLICATION NUMBER: 08/979,673  
 EARLIER FILING DATE: 1997-11-26  
 EARLIER APPLICATION NUMBER: 09/107,657  
 EARLIER FILING DATE: 1998-06-30  
 NUMBER OF SEQ ID NOS: 34  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 34  
 LENGTH: 618  
 TYPE: PRT  
 ORGANISM: Talaromyces emersonii  
 US-09-199-290-34

Query Match 61.0%; Score 1689; DB 3; Length 618;  
 Best Local Similarity 61.2%; Pred. No. 9,1e-138;

```
Matches 322; Conservative 75; Mismatches 111; Indels 18; Gaps 5;
QY 7 LALSGLVCTGLANV--ISKRAI--LDSWLSNEATVARTALINNIGADGAWYSGADSGIVVA 63
Db 9 LCIIGLTPAFAFAPVABARATGSLDSFLATEETPIALOGVANNIGPNAGADVAGASAGIVVA 68
QY 64 SPSTDNDEYFTYTRDGLVLTIVDLFRNGDTSLSLTIENTYISAQIVOGISNPSGDL 123
Db 69 SPSPSDNPNYFTSWTRDALTRAKYLVDFAFIAENKPLEOTIQOYISAQKVQTIISNPSGDL 128
QY 124 SGAGLGPKNVDEATYATGSGWRPQRDGPALRATAMIGFQWMLDNGSTPATDIWPLV 183
Db 129 TG-GLGPKRVNVEATATGPGWRPQRDGPALRATALIAVANYLIDNGEASTADEIIPVIV 187
QY 184 RNDLSYAOYWNQGYDLMEEVNNGSFFTTAVOHRALVEGSAFATAVSSCSMCDQAPE 243
Db 188 QNDLSYITQYWNSTFPLMEVEGSSFFTTAVOHRALVEGNALATRLNHTCSNCSQAPE 247
QY 244 ILCTLOSFMTGSFLANF--DSRSRGKDANTLGSIHFPDEPAACDDSTFQPCSPRALANH 302
Db 248 VLCTLOSYMTGSVYLANFGSGRGKQVNSILGSIHFPDPAAGCDDSTFQPCSPRALANH 307
QY 303 KEVVDSPRSIYTLNDGLSDSEAVAVAGRYPEPTYNGNPMFLCTLAABOLYDALYQMDKQ 362
Db 308 KVTYDSFRSIYALNSGIAESGSAVAVGRYPEDVYQGNPMFLATATAAEOLYDALYQMKI 367
QY 363 GSLEVTDSLDFPKALYSDAATGYSSSSSTYSISYDAVATPADGFSIYETHAASNGSM 422
Db 368 GSISITDVSLEPFQDIPSAAVGTYNNGSTTFNDIISAVOYTGDIYIEKTYTPSDGSL 427
QY 423 SEQYDKSDGEOLASRDITWMSYALLTANNRNSVVPASWGETSASSVPCTCAATSAGTY 482
Db 428 TEQFSRDTGPIASALTWASALITASARQSVVPASWGESASSVPACSAFSAATG 487
QY 483 SSVTVTSMPSIVATGTTTATPTGSGSVTSKTTATASKTSTT 528
Db 488 STATNTWPS-----SGSGSSTTSSAPCTTPTISAVT 520

RESULT 10
US-09-821-616-34
; Sequence 34, Application US/09821616
; Patent No. 6620924
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Ruby
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279.200-US
; CURRENT APPLICATION NUMBER: US/09/821,616
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/199,290
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1557/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 0925/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/070,746
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/094,344
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/979,673
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/107,657
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Talaromyces emersonii
US-09-821-616-34
```

```
Query Match 61.0%; Score 1689; DB 4; Length 618;
Best Local Similarity 61.2%; Pred. No. 9.1e-138;
Matches 322; Conservative 75; Mismatches 111; Indels 18; Gaps 5;
QY 7 LALSGLVCTGLANV--ISKRAI--LDSWLSNEATVARTALINNIGADGAWYSGADSGIVVA 63
Db 9 LCIIGLTPAFAFAPVABARATGSLDSFLATEETPIALOGVANNIGPNAGADVAGASAGIVVA 68
QY 64 SPSTDNDEYFTYTRDGLVLTIVDLFRNGDTSLSLTIENTYISAQIVOGISNPSGDL 123
Db 69 SPSPSDNPNYFTSWTRDALTRAKYLVDFAFIAENKPLEOTIQOYISAQKVQTIISNPSGDL 128
QY 124 SGAGLGPKNVDEATYATGSGWRPQRDGPALRATAMIGFQWMLDNGSTPATDIWPLV 183
Db 129 TG-GLGPKRVNVEATATGPGWRPQRDGPALRATALIAVANYLIDNGEASTADEIIPVIV 187
QY 184 RNDLSYAOYWNQGYDLMEEVNNGSFFTTAVOHRALVEGSAFATAVSSCSMCDQAPE 243
Db 188 QNDLSYITQYWNSTFPLMEVEGSSFFTTAVOHRALVEGNALATRLNHTCSNCSQAPE 247
QY 244 ILCTLOSFMTGSFLANF--DSRSRGKDANTLGSIHFPDEPAACDDSTFQPCSPRALANH 302
Db 248 VLCTLOSYMTGSVYLANFGSGRGKQVNSILGSIHFPDPAAGCDDSTFQPCSPRALANH 307
QY 303 KEVVDSPRSIYTLNDGLSDSEAVAVAGRYPEPTYNGNPMFLCTLAABOLYDALYQMDKQ 362
Db 308 KVTYDSFRSIYALNSGIAESGSAVAVGRYPEDVYQGNPMFLATATAAEOLYDALYQMKI 367
QY 363 GSLEVTDSLDFPKALYSDAATGYSSSSSTYSISYDAVATPADGFSIYETHAASNGSM 422
Db 368 GSISITDVSLEPFQDIPSAAVGTYNNGSTTFNDIISAVOYTGDIYIEKTYTPSDGSL 427
QY 423 SEQYDKSDGEOLASRDITWMSYALLTANNRNSVVPASWGETSASSVPCTCAATSAGTY 482
Db 428 TEQFSRDTGPIASALTWASALITASARQSVVPASWGESASSVPACSAFSAATG 487
QY 483 SSVTVTSMPSIVATGTTTATPTGSGSVTSKTTATASKTSTT 528
Db 488 STATNTWPS-----SGSGSSTTSSAPCTTPTISAVT 520

RESULT 11
US-09-199-290-7
; Sequence 7, Application US/09199290
; Patent No. 6255084
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Ruby
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279.200-US
; CURRENT APPLICATION NUMBER: US/09/199,290
; CURRENT FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1557/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 0925/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/070,746
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/094,344
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/979,673
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/107,657
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Talaromyces emersonii
US-09-199-290-7
```





```

Db      66 ASGIIVASPDKTNBDYWTYTRDSALTFKCVDTFTNSYDASLQAEIONTYIVAAQHLQGV 125
Qy      116 SNPSGDLSSGAGLGEPEKFNVDETAYTGSWGRPORBGPALRATAMIGFQWLLDNGYSTA 175
      126 SNPSGSLSDSGLGEPEKFNVDMSOFTGAMGRPOBGPALRATAMIGFQWLLDNGYSTA 185
Qy      176 TDIVWPLVRNDLSYVAQYWNQGYDLMEVNGSSFFTIYAVQHRALVEGSAFATAVGSSCS 235
      186 SSIWVPYIKDILAVYAO--NNTGFDLMEVNGSSFFTIYAVQHRALVEGSAFATAVGSSCS 243
Qy      236 WCOSQAEIILCYLOSFTGSG--FILLNPDSSRSKQDANTLLGSIHTDPEAACDSTFOP 293
      244 ACSVAQAQILCFQGSFWSGSSGYILLAN---STAKDANTLLGSIHTDPEAACDSTFOP 299
Qy      294 CSPPALANKEVDSFPRSIYTLNDGLSDSEAVAGRYPEPTYNGNPMFLCTLLAAEQLY 353
      300 CSDPALANHKVYTDAPFSIYSINGIAEGSAVANGRPEDSPFGNPMYINTLLAAEQLY 359
Db      354 DALYQMDKQSLVETDVSLEDFKALYSDAATGYSSSSSTYSIVDAVKTFADGFSVIVE 413
      360 DALYVWKQGSITVTSTSLAFKDFSSITPGYSSSTSTYTTLYNNAISAVADGYNMIVA 419
Qy      414 THASNSMSRQYKSGEQLSARDLTWSYVALLTANNRNSVVPASMGETSASVPGTC 473
      420 QYAGTNSLSBQFCKTGEPLSAYDLTWSYVALLTAAARRAGVPPSMAASANSVPAQC 479
Qy      474 AATSAIGTYSVYVTSWPSIVATGTTTATPTGSGSVTSTKTTATASKTSTT 528
      480 SATSVGSYSATATSTFP-----SQTPRASSTAGS-SPASSTTATATACSTT 527

```

RESULT 14  
US-08-596-300A-7  
Sequence 7, Application US/08596300A  
Patent No. 5834191

## GENERAL INFORMATION:

## APPLICANT:

TITLE OF INVENTION: Production of Heterologous Peptides

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber &amp; Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/596,300A

FILING DATE: 13-FEB-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1321-1-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 626 amino acids

TYPE: amino acid

STRANDBONES: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

FRAGMENT TYPE:

US-08-596-300A-7

Query Match 56.9%; Score 1577; DB 2; Length 626;  
Best Local Similarity 57.6%; Pred. No. 4.8e-128;  
Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

```

Qy      2 SPSLALSLGLVCTGLANVISKRAITLDSMLSNETAVKATIIANNIGADGAWGADSGIV 61
      13 AFAVAVGLPPLLEKRRSDIKR-SVDSYIQTEPTPLAQKNLLCNIGSGRAGAGAGV 71
Db      62 VASPTDNPDYFYTWRTDGLVLTVDLPRNG-DTSLSTIENIENISAQAVOGISPPSG 120
      72 VASPSKSPDYWTWTRDALVTKLIVDEFTDNYNTLQNTIQAIAAQAQKQGVNSPSG 131
Qy      121 DLSGAGLGEPEKFNVDETAYTGSWGRPORBGPALRATAMIGFQWLLDNGYSTADIVW 180
      132 SLNSGAGLGEPEKFNVDLQFTGAMGRPORBGPALRATAMIGFQWLLDNGYSTADIVW 191
Db      181 PLVRNDLSYVAQYWNQGYDLMEVNGSSFFTIYAVQHRALVEGSAFATAVGSSCSWDQ 240
      192 PIVYNDLAYTAQYWNMGFPLMEVNGSSFFTIYASHRALVEGSAFATAVGSSCSACDAI 251
Qy      241 APRLICYLOSFTGSG-FILLNPDSSRSKQDANTLLGSIHTDPEAACDSTFOPCSPRAL 299
      252 APQLICFQGSFWSGSSGYILLAN---STAKDANTLLGSIHTDPEAACDSTFOPCSPRAL 311
Db      300 ANHKEVDSFPRSIYTLNDGLSDSEAVAGRYPEPTYNGNPMFLCTLLAAEQLYDALYQW 359
      312 ANHKVVDMSR-FWGVNSGRTAGKAAVGRYVADVYNGNPMYINTLLAAEQLYDAYVW 370
Qy      360 DKQGSLEVTVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTFADGFSIYETHAASN 419
      371 KQGSITVTSTSLAFKDFSSITPGYSSSTSTYTTLYNNAISAVADGYNMIVAQYVPSD 430
Db      420 GMSSEQYDKSDGQLSARDLTWSYVALLTANNRNSVVPASMGETSASVPGCAATSAL 479
      431 GSLAQFDKSGAPLSYTHLTWSYVALLTAAARRAGVPPSMAASANSVPAQCASATVA 490
Qy      480 GTYSVTVTSWPSIVATGTTT-ATPTGSGS-----VTSTSKTTATASKT 524
      491 GSYATATATSTFPANLTPASSTTVPPTQTCADHVELVTNEKVTTSYGCT 541

```

RESULT 15  
US-08-596-300A-14  
Sequence 14, Application US/08596300A  
Patent No. 5834191

## GENERAL INFORMATION:

## APPLICANT:

TITLE OF INVENTION: Production of Heterologous Peptides

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber &amp; Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/596,300A

FILING DATE: 13-FEB-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1321-1-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684



TELEX: 133521  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 626 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 FRAGMENT TYPE: NO  
 US-08-596-300A-14

Query Match 56.9%; Score 1577; DB 2; Length 626;  
 Best Local Similarity 57.6%; Pred. No. 4,8e-128;  
 Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

QY 2 SFRSLALSGVCTGLANVISKRALDSWLSNEATVARTAIINIGADGAWVGADSGIV 61  
 DB 13 AFOAVTGLPDPHEKSHDIIR-SVDSYIQETETPIAQCNILCNIGASGCRASGAGV 71  
 QY 62 VASPESTDNDPYFTYTRDSGLVTKLVDFRNG-DTSLSTIENTISAQAIVOGISNPSG 120  
 DB 72 VASPEKSSPDYWTYTRDALVTKLVDEFINDYNTLQNTIQAYAAQAQKQGVSNPSG 131  
 QY 121 DLSGAGLGEPEKFNDEFAVTSWGRPORDBPALRATAMIGFGOWLNDNGYSTATDIW 180  
 DB 132 SLSNAGLGEPEKFNVDLQOFTGAKGRFQDGPPLRAIALIGKMLVSNNGYADTKSIW 191  
 QY 181 PLVRNDLSVAQYMNQGYDLWEVNGSSFTTIAVOHRALVEGSAFATAVGSSCSMCD 240  
 DB 192 PIVKNDLAYTAQYMNNTGFDLWEVNSSSFTTIAASHRALVEGSAFAKSVGSSCSACDAI 251  
 QY 241 APELCYLOSFWTGS-FILANFDSRSRGKDANTLLGSIHTPEPAACDSTFQCSPRAL 299  
 DB 252 APOILCFQOOSFWSNSGYIISNFVNYRSGKDINSVLTSHNFPAGCDVNTFQPCSDRAL 311  
 QY 300 ANHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPFELCTIAAEQLYDALYOW 359  
 DB 312 ANHKYVVDMSR-FMGVNSGRTAGKAAVGRYAEVDYINGNPMYLAITLAAAEQLYDAVYW 370  
 QY 360 DKQGSLEYTVDSLDFEKLALYSDAATGYSSSSSTYSIVDAVKTFAFGFVSIVETHAASN 419  
 DB 371 KKQGSITVTSTLAFKDLVPSVSTGYSSSSSTYTAIINAVTYADGFVDIVAQYTPSD 430  
 QY 420 GSMSEQYKSDGEQLSARDLTSYALALTANNRNSVVPASWGETSASVPGTCAATSAI 479  
 DB 431 GSLAEQFPKDSGAPLSATHLTWSYASFLSAAARRAGIIVPPSWGAASANSPLPGSCASTVA 490  
 QY 480 GTYSSTVTWSPSIVATGTTT-ATPTGSSG-----VTSTSKTATATASKT 524  
 DB 491 GSYATATATSPANLTPASTVTPPTQTCGCAADHEVLTFENEKVTTSYGOT 541

Search completed: June 17, 2004, 17:48:36  
 Job time : 25 secs

**HIS PAGE LEFT BLANK**



DR InterPro: IPR000165; Glyco\_hydro\_15.  
DR InterPro: IPR008928; Glyco\_trans\_6hp.  
DR Pfam: PF00686; CBM\_20; 1.  
DR Pfam: PF00723; Glyco\_hydro\_15; 1.  
DR PRINTS: PR00736; GLHYDRASE15.  
DR ProDom: PD001568; CBD 4; 1.  
DR PIRSF: PIRSF01031; GLY-a-glycd SBD; 1.  
SQ SEQUENCE 639 AA; 68308 MW; 02ADB56B6542B69 CAC64

Query Match	96.3%	Score 2669.5;	DB 3;	Length 639;
Best Local Similarity	98.1%;	Pred. No. 3.9e-162;		
Matches 521;	Conservative 1;	Mismatches 6;	Indels 3;	Gaps 3;

Qy	61	VVASPESTNDNDYYTWTTPRSGVLKTVLFPNGGTSLLSTLENTISAQALVQGISNPSG	120
Db	61	VVASPESTNDNDYYTWTTRDSGLVKTVLDFRNGGTSLLSTLENTISAQALVQGISNPSG	120
Qy	121	DLSSGAGLGEPKFNVDETVAYTGSWGRPQRDGALPATAMI GFGQMLDNGYSTATDIW	180
Db	121	DLSSGAGLGEPKFNVDETVAYTGSWGRPQRDGALPATAMI GFGQMLDNGYSTATDIW	180
Qy	181	PLVRNLSVYAQVYNQTVGLMEVYNGSSFITANQHALVEGSAFATAYSSSCWCDQ	240
Db	181	PLVRNLDLSVYAQVYNQTVGLDM-EVANGSSFITAAQHALVEGSAFAFAYSSSCWCDQ	239
Qy	241	APRLICYLGSFWMGSEFLANFDSRSRSGKANT-LVGSIHTEFDPKAACDSTFQCSRAL	299
Db	240	APRLICYLGSFWMGSEFLANFDSRSRAKQANLTLGSIHTTDPKAACDSTFQCSRAL	299
Qy	300	ANRKEVYDSEFRS1YTLNDGLSDSBAVAGRYPEDTYNQNPMFLCTLAAEQLYDALYQW	359
Db	300	ANRKEVYDSEFRS1YTLNDGLSDSEAVAVAGRYPEDTYNQNPMFLCTLAAEQLDALYQW	359
Qy	360	DKGGSLEVTDSLDFEFKALYSDAAGTYSSSSSSTYS1VDVAKTPADGPFYSIVETHAASN	419
Db	360	DKGGSLEVTDSLDFEFKALYSD-AGTYSSSSSSTYS1VDAAKTPADGPFYSIVETHAASN	418
Qy	420	GSMSEQYDSDSGEOLSARDLITWSYALLTANNRRNSVVPASMGETSAS5VGTCAATSAI	479
Db	419	GSMSEQYDSDSGEOLSARDLITWSYALLLTANNRRNVVPASMGETSAS5VGTCAATSAI	478
Qy	480	GYTSVTVYVSWPSIVATGGTTTAAFTPTGSGSYTSRSKTTAATAKSTSTTNS	530
Db	479	GYTSVTVYVSWPSIVATGGTTTAAFTPTGSGSYTSRSKTTAATAKSTSTTNS	529

ID	Q12537	PRELIMINARY;	PRT;	639 AA.
AC	Q12537/			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, last annotation update)			
DE	Glucanamyase precursor (EC 3.2.1.3).			
GN	GLA.			
OS	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus			
OX	NCBI_TaxId=105351;			
RN	[1] _			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=X-100;			
RA	Diagne I., Pechexonov V.T., Bulat S.A., Firsov I.M.;			
RT	"A glucanamyase gene from Aspergillus awamori X-100: structure,			
RT	allocation, and gene phylogeny.";			
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; U93103; AAB02927.1; -.			
DR	HSSP; P04064; 1GAI.			
DR	GO; GO:0005509; F:calcium ion binding; IEA.			

DR	GO:	GO:0004333;	F:glucan 1,4-alpha-glucosidase activity; IEA.
DR	GO:	GO:0016788;	F:hydrolyase activity, acting on glycosyl bonds; IEA.
DR	GO:	GO:0005976;	P:polysaccharide metabolism; IEA.
DR	InterPro:	IPIR002044;	CBD_4.
DR	InterPro:	IPIR002048;	E-hand.
DR	InterPro:	IPIR008291;	Glu-a-glycd_SBD.
DR	InterPro:	IPIR000165;	Glyco_hydro_15.
DR	InterPro:	IPIR008928;	Glyco_trans_6hp.
DR	Pfam:	PF00666;	CBM_20; 1.
DR	Pfam:	PF00723;	Glyco_hydro_15; 1.
DR	PRINTS:	PR00736;	GLHYDRLASE15.
DR	ProDom:	PD001568;	CBD_4; 1.
DR	PROSITE:	PS00018;	E_HAND; 1.
DR	PROSITE:	PS00820;	GLUCOMYLASE; 1.
DR	PIRSF:	PIRSF001031;	Glu-a-glycd_SBD; 1.
KW	Glycosidase; Hydrolase; Signal.		
FT	SIGNAL	1	POTENTIAL.
ET	CHAIN	25	GLUCOMYLASE.
SO	SEQUENCE	639 AA;	68277 MW; 6E993DD637D174ACB CRC64;

Query Match	92.5%	Score 2564.5;	DB 3;	Length 639;
Best Local Similarity	92.3%;	Pred. No. 2e-15;		
Matches 494;	Conservative 16;	Mismatches 14;	Indels 11;	Gaps 3;

Qy	MsRSRLtASGLVCtCAVtKtANvISKRAtLDtSMtSNEtVtARtILNNtGADtAMtSGAdSGt	60
Db	1 MsRSRLtASGLVCtCSlASvISKRAtLDtSMtSNEtVtARtILNNtGADtAMtSGAdSGt	60
Qy	61 VVASpSTNDPdyFYtWTNRDStGLvLkTlYdLFRNGDStLSStENyISAOAtVOCISnPSG	120
Db	61 VVASpSTNDPdyFYtWTNRDStGLvLkTlYdLFRNGDStLSStENyISAOAtIOGVnPSG	120
Qy	121 DLSsGtGLsPpKFNVDtEYtGtSMGRpORDGpALtATAmIcFGOMLdNGtStATdIYw	180
Db	121 DLSsG--GLsPpKFNVDtEYtGtSMGRpORDGpALtATAmIcFGOMLdNGtStATdIYw	179
Qy	181 pLVrNDLStVAOYwNQtGYdLwEEvNGSSpFTtAVOHrALvESGAfAtAVAGSScSwCDsQ	240
Db	180 pLVrNDLStVAOYwNQtGYdLwEEvNGSSpFTtAVOHrALvESGAfAtAVAGSScSwCDsQ	239
Qy	241 ApRILtCYLdSf-----KtGSfFLANpDSSRGKAntLlGSIHfFDpEACdDStFQPCs	295
Db	240 ApILtCYLdSfSGSRSGwtGSYtLANpD-----KdNtLlGSIHfFDpEACdDStFQPCs	294
Qy	296 pRALANhKEvDvSFRStYtLNDGLdSdSAvAVAGrYpEDtYvNGNpWfLCTlAAAEOLYdA	355
Db	295 pRALANhKEvDvSFRStYtLNDGLdSdSAvAVAGrYpEDtSYvNGNpWfLCTlAAAEOLYdA	354
Qy	356 LYOMdKQSLtETDvSLDpFFKALySDAtNGYSSSSStSYsIYNAtVtPDpGfSYtETH	415
Db	355 LYOMdKQSLtETDvSLDpFFKALySGAAtNGYSSSSStSYsIvAAtVtPDpGfSYtETH	414
Qy	416 AASNGSMsEQYKSDOEQLsARDLtwSYAALtANNRNSvVAPASMEtSASvPGtCAA	475
Db	415 AASNGSLsBQPKSDODELSARDLtwSYAALtANNRNSvVAPsMEtSASvPGtCAA	474
Qy	476 tSAIGtYSSvYtYtSPMSIvAtNGtTtATpFGGSsYtStKtTAtTAKtStTtRS	530
Db	475 tSASGtYSSvYtYtSPMSIvAtNGAtTtATtTtGGSgGvYtStKtYtTAtTAKtStTtRS	529

RESULT	3
09C1V4	
ID	09C1V4
AC	09C1V4; PRELIMINARY; PRT; 618 AA.
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE	Glucosylase precursor (EC 3.2.1.3).
GN	GA.
OS	Talaromyces emersonii.
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC	Eurotiales; Trichocomycetes; Talaromyces.

OX NCBI\_TaxID=68825;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nielsen B.R., Lembeck J., Frandsen T.P.;  
 RT "Cloning, heterologous expression, and enzymatic characterization of a  
 RT thermostable glucosylase from *Talaromyces emersonii*,"  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ304803; CAC28076.1; -  
 DR HSSP: P04064; 1GAT.  
 DR GO: GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.  
 DR GO: GO:0016798; F:hydrolyase activity, acting on glycosyl bonds; IEA.  
 DR GO: GO:0005976; P:polysaccharide metabolism; IEA.  
 DR InterPro: IPR002044; CBD 4.  
 DR InterPro: IPR008291; Glu-a-glycd SBD.  
 DR InterPro: IPR00165; Glyco\_hydro\_15.  
 DR InterPro: IPR008928; Glyco\_trans\_6np.  
 DR Pfam: PF00686; CBM\_20; 1.  
 DR Pfam: PF00723; Glyco\_hydro\_15; 1.  
 DR PRINTS: PR00736; GLHYDRLASE15.  
 DR ProDom: PD001568; CBD 4; 1.  
 DR PIRSF: PIRSF00103; Glu-a-glycd\_SBD; 1.  
 DR GlycoStase; Hydrolase; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 28 618 GLUCOMYLASE.  
 SQ SEQUENCE 618 AA; 65429 MW; C48A034A2C06E5B2 CRC64;  
 Query Match 61.0%; Score 1689; DB 3; Length 618;  
 Best Local Similarity 61.2%; Pred. No. 1,4e-99;  
 Matches 322; Conservative 75; Mismatches 111; Indels 18; Gaps 5;  
 Oy 7 LALSGVCTGLANV-ISKRAAT--LDLSWLSNEATVARTALINNGADGAMVSGDSIVYA 63  
 Db 9 LCILGTPAAPAPAPAPAPAPATSSLSDFLATEPTIALQGVLNINIGPAGVAGASGIVYA 68  
 Oy 64 SESTNDPDYFYTWTRDSGLVLTVDLFRNGDTSLSSTIENYISAQAVQGISNPGDLS 123  
 Db 69 SPSRSDDPNFYFWTRDAALTAKYLVDAFAGNKMLEGTIQITQISQAQKQITISNPGDLS 128  
 Oy 124 SGAGLGEPEFNVDETAATYTSKGRPQRDGPALRATAMIGCGWLLDNGYSTATDIWPLV 183  
 Db 129 TG-GGEPKFNVENETAFTGPMGRPQRDGPALRATAMIAVANVLDNGEASTDEIWPV 187  
 Oy 184 RNDLSVYAOYNNQGTGDLMEEVNGSSFFITAVOHRALVEGSAFATVAGSSGCSQAPET 243  
 Db 188 QNDLSYITQYNNSSFTDLMEEVGSSFFITAVOHRALVEGSAFATVAGSSGCSQAPET 247  
 Oy 244 ILCYLQSFWTSGFIILANF-DSSRSGKDANTLLGSIHTFPBEACDSTFQPCSPRALANH 302  
 Db 248 VLCFLQSYWTSGLVLANFGSGRSGKDVNSILGSIHTFPDAGCDDSTFQPCSPRALANH 307  
 Oy 303 KEVDSFRSITYTLNDGLDSEAVANGRYPEPTYNGNPMFLCTLAAEQLYALYQMDKQ 362  
 Db 308 KVVTSFRSIVAINSGIAEGSAVANGRYPEPTYNGNPMFLCTLAAEQLYALYQMDKQ 367  
 Oy 363 GSELETVDSLDFPKKALYSDAATGTYSSSSSTVSDVADKTPADGFSIVETHAASNGSM 422  
 Db 368 GSISTIDVSLDFPKKALYSDAATGTYSSSSSTVSDVADKTPADGFSIVETHAASNGSM 427  
 Oy 423 SEQYKSDGEOLASARDLTWSYAALLTANNRRNSVVPASGSETASVPQCATSAIGTY 482  
 Db 428 TEQFRTDGTPLASALTLTWSYASLLTASARRQSVVPASGESSASVPVCAITATGY 487  
 Oy 483 SSVYTTWSPSIVATGTTTATPTGSGVTSSTKTTATASKTSITT 528  
 Db 488 STATNTWSPS-----SSGSGSTTSSAPCTTPTSAVT 520

RESULT 4  
 ID 059846 PRELIMINARY; PRT; 493 AA.  
 AC 059846;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Glucosylase.  
 OS Aspergillus oryzae.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eucotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 OC NCBI\_TaxID=5062;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O-1013;  
 RX MEDLINE=9817274; PubMed=9511753;  
 RA Hata Y., Ishida H., Ichikawa E., Kawato A., Suginami K., Imaiya S.;  
 RT "Nucleotide sequence of an alternative glucosylase-encoding gene  
 RT (glab) expressed in solid-state culture of *Aspergillus oryzae*,"  
 RL Gene 207:127-134(1998).  
 DR HSSP: P04064; 1GAT.  
 DR PIR: PC6503; JC6538.  
 DR GO: GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.  
 DR GO: GO:0005976; P:polysaccharide metabolism; IEA.  
 DR InterPro: IPR00165; Glyco\_hydro\_15.  
 DR InterPro: IPR008928; Glyco\_trans\_6np.  
 DR Pfam: PF00723; Glyco\_hydro\_15; 1.  
 DR PRINTS: PR00736; GLHYDRLASE15.  
 DR PROSITE: PS00820; GLUCOMYLASE; 1.  
 SQ SEQUENCE 493 AA; 52401 MW; 3707B3A6F0C66256 CRC64;  
 Query Match 54.8%; Score 1518; DB 3; Length 493;  
 Best Local Similarity 57.5%; Pred. No. 8,6e-89;  
 Matches 280; Conservative 82; Mismatches 123; Indels 2; Gaps 2;  
 Oy 5 SLIALSGVCTGLANV-ISKRAATLDSWLSNEATVARTALINNGADGAMVSGDSIVYA 64  
 Db 8 SLIALAGVAPHPSPFIHRQSDLNAFIAQPTIAKQGVLNINIGADKIVEGAAGIVVA 67  
 Oy 65 PSTNDPDYFYTWTRDSGLVLTVDLFRNGDTSLSSTIENYISAQAVQGISNPGDLS 124  
 Db 68 PSKSNPDYFYTWTRDAGLMEVLEQFTGGDATTSTIQNVYDSQANDQAVSNPGGLSD 127  
 Oy 125 GAGLGEPEFNVDETAATYTSKGRPQRDGPALRATAMIGCGWLLDNGYSTATDIWPLV 184  
 Db 128 GSGLAEPFVYNNISGFTSWKGRPQRDGPALRATAMIAVGNLSISDQSVVANNIWPVQ 187  
 Oy 185 NDLSTVYAOYNNQGTGDLMEEVNGSSFFITAVOHRALVEGSAFATVAGSSGCSQAPET 244  
 Db 188 NDLSTVYAOYNNQGTGDLMEEVNGSSFFITAVOHRALVEGSAFATVAGSSGCSQAPET 246  
 Oy 245 LCVLQSFWTSGFIILANFSS-RSGKDANTLLGSIHTFPBEACDSTFQPCSPRALANH 303  
 Db 247 LCHLQDFPNNGAVSNLPTNGSGSLDNTLSLGSHTFPDPAACDSTFQPCSPRALANH 306  
 Oy 304 EYVDSFRSITYTLNDGLDSEAVANGRYPEPTYNGNPMFLCTLAAEQLYALYQMDKQ 363  
 Db 307 LVVDSFRSIVYGINNRGAKAAAVGPYAEPTYNGNPMFLCTLAAEQLYALYQMDKQ 366  
 Oy 364 SLEVTDSLDFPKKALYSDAATGTYSSSSSTVSDVADKTPADGFSIVETHAASNGSM 423  
 Db 367 QNVVTEISLDFPKKALYSDAATGTYSSSSSTVSDVADKTPADGFSIVETHAASNGSM 426  
 Oy 424 BOYKSDGEOLASARDLTWSYAALLTANNRRNSVVPASGSETASVPQCATSAIGTY 483  
 Db 427 EYRSRQGTTPVASDLTWSYAALLTANNRRNSVVPASGSETASVPQCATSAIGTY 486  
 Oy 484 SVYTTWSP 490  
 Db 487 TPTVGSW 493

## RESULT 5

012623 PRELIMINARY; PRT; 620 AA.  
 ID 012623  
 AC 012623;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Glucoamylase (EC 3.2.1.3).  
 GN GLAI.  
 OS Humicola grisea var. thermoides.  
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.  
 OX NCBI\_TaxID=5528;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Betka R.M., Rey M.W., Thompson S.A., Gray G.L., Carmona C.L.,  
 RA Power S.D.;  
 RT "Molecular cloning, analysis and expression of the glai gene encoding  
 RT a thermostable, raw starch-digesting glucoamylase from the fungus  
 RT Humicola grisea var. thermoides";  
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M89475; AAA3386.1; -.  
 DR HSSP; P04064; IGAI.  
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0005976; P:polysaccharide metabolism; IEA.  
 DR InterPro; IPR002044; CBD\_4.  
 DR InterPro; IPR008291; Glu-a-glycosid\_SBD.  
 DR InterPro; IPR00165; Glyco\_hydro\_15.  
 DR InterPro; IPR008928; Glyco\_trans\_6hp.  
 DR Pfam; PF00686; CBM\_20; 1.  
 DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
 DR PRINTS; PR00736; GLHYDRLASE15.  
 DR PRODOM; PD001568; CBD\_4; 1.  
 DR PROSITE; PS00820; GLUCOAMYLASE; 1.  
 DR PIRSF; PIRSF001031; Glu-a-glycosid\_SBD; 1.  
 DR GlycoSite; Hydrolase.  
 KM SEQUENCE 620 AA; 66525 MW; 54214FF67E20BE0A CRC64;  
 SQ

Query Match 50.5%; Score 1400; DB 3; Length 620;  
 Best Local Similarity 49.5%; Pred. No. 4e-81;  
 Matches 269; Conservative 96; Mismatches 146; Indels 32; Gaps 7;  
 QY 2 SFRSLALSGVTCGL-----ANVTSKRATLDSWLSNEATVARTAILNINIGADGAWVSGA 56  
 DB 3 TFSKLVLVGSVAGSALGRPHSSSRKQERAAVDITINTKPLAMKNLANNIPNGKAAPGA 62  
 QY 57 DSGIVVASPTDNDPYFYTWTRDSGLVLTVD-LFRNGDTSLSLTIENTYISAQAIYOGI 115  
 DB 63 AAGVVASPSRTDPPYFTWTPDALVLTGIIESLGHNYNTTL-----QQV 108  
 QY 116 SNBEGDLSGAGLGEPEKNVDETATYTSWGRPPORDGPARLATAMIGFQWLLDNGYSTA 175  
 DB 109 SNPSGTADGSGLGEAKFNVDLTAFETGWRGPRDGPRLAIALIQYAKWLIANGYSTA 168  
 QY 176 TDIWPLVRLNDLSVYAQYNQOTGYDLMEEVNNGSFFTTIAVOHRLVSGAATATVAGSSCS 235  
 DB 169 KSVWMPVVKNDLATTQAQWNETGFDLMEVPGSSFFTTIASHRALTBGAYILAQLDTECP 228  
 QY 236 WCDQABEILCYLQSFWT--GSFILANFDS--RSGKDANTLLGSINTFDEEACDSTF 291  
 DB 229 PCTTVAPOVLCFQQAFAFNKSNVYVSTAGEYRSGKANSILASIHNFDEACDNLTF 288  
 QY 292 QPCSPRALANKEVVDVFRSLYTLNDGLSDEANAVGRIPEDTTYNGNPMFLCTLAAEQ 351  
 DB 289 QPCSPRALANKEVVDVFRSLYTLNDGLSDEANAVGRIPEDTTYNGNPMFLCTLAAEQ 348  
 QY 352 LYDLAYOMDKGSLVETDVSLLDFKALYSDAATGYSSSSSTYSIVAYVTPADGFSI 411  
 DB 349 LYDLAYWNNKGSITVIVSVLPFRFDLVSVSTGYTSSSSTFTNIVAVKAVADGFLFV 408  
 QY 412 VETHAASNGSMSEQYDKSDGQSLARDLTWSVAAILLTANNRRNSVVPASW-GETSASSVP 470

DB 409 AAKYTPSNGALAEQYDNTGKPDSDADLTWVSFAFLSALDRRAGLVPPSRASVAKSOLP 468  
 QY 471 GTCAATATIGTGSVTVTWSPIVATGCTTTTATPSSGVSSTSKTATATASKSTTRS 530  
 DB 469 STGSRLEVAGTVAAISTSPS-----KOTPNBSAAPSPPYPTACADASEVYTFNE 521  
 QY 531 GMS 533  
 DB 522 RVS 524

## RESULT 6

012596 PRELIMINARY; PRT; 579 AA.  
 ID 012596  
 AC 012596;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Glucoamylase G2 (EC 3.2.1.3).  
 OS Corticium rolfsii.  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 OC Aphyllophorales; Corticiaceae; Corticium.  
 OX NCBI\_TaxID=39291;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=AHU9627;  
 RX MEDLINE=96158471; PubMed=8597548;  
 RA Nagasaka Y., Muraki N., Kimura A., Suto M., Yokota A., Tomita F.;  
 RT "Cloning of Corticium rolfsii glucoamylase cDNA and its expression in  
 RT Saccharomyces cerevisiae";  
 RL Appl. Microbiol. Biotechnol. 44:451-458(1995).  
 DR EMBL; D49448; BAA08436.1; -.  
 DR HSSP; P04064; IGAI.  
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0005976; P:polysaccharide metabolism; IEA.  
 DR InterPro; IPR002044; CBD\_4.  
 DR InterPro; IPR008291; Glu-a-glycosid\_SBD.  
 DR InterPro; IPR00165; Glyco\_hydro\_15.  
 DR InterPro; IPR008928; Glyco\_trans\_6hp.  
 DR Pfam; PF00686; CBM\_20; 1.  
 DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
 DR PRINTS; PR00736; GLHYDRLASE15.  
 DR PRODOM; PD001568; CBD\_4; 1.  
 DR PIRSF; PIRSF001031; Glu-a-glycosid\_SBD; 1.  
 DR GlycoSite; Hydrolase.  
 KM SEQUENCE 579 AA; 61548 MW; FF56D83F9D1A9BC9 CRC64;  
 SQ

Query Match 50.3%; Score 1394; DB 3; Length 579;  
 Best Local Similarity 51.2%; Pred. No. 8.8e-81;  
 Matches 294; Conservative 77; Mismatches 151; Indels 52; Gaps 11;  
 QY 3 FRSLLALSGVTCGLANVTSKRATLDSWLSNEATVARTAILNINIGADGAWVSGDGLIV 62  
 DB 2 FRSLLALA---ACAVASASQASATAYLTRESAVAKNGVLCNIGSGCSGSEAGSIV 58  
 QY 63 ASPTDNDPYFYTWTRDSGLVLTVDLFRNG-DTSLSTIENTYISAQAIYOGISNPSGD 121  
 DB 59 ASPTKISPDVLTWTRDSLSLFFKMLIDQYNTGDLTTLRLTLDVSVSEATITIQTSNSSGT 118  
 QY 122 LSSGAGLGEPEKNVDETATYTSWGRPPORDGPARLATAMIGFQWLLDNGYSTATDIWMP 181  
 DB 119 VSTG-GLGEPEKNVDETATYTSWGRPPORDGPARLATAMIGFQWLLDNGYSTATDIWMP 177  
 QY 182 LVRLNDLSVYAQYNQOTGYDLMEEVNNGSFFTTIAVOHRLVSGAFAFATVAG--SSCSKCD 239  
 DB 178 IIRKLDLYVNSDMNQITFDLMEEVDSSTFTTAVOHRALVQGAFAFATLICQTSASTYSA 237  
 QY 240 QAPBILCYLQSF--TGSFILANFDS--RSGKDANTLLGSINTFDEEACDSTPOPCSPRA 298  
 DB 238 TAPBILCYLQSF--TGSFILANFDS--RSGKDANTLLGSINTFDEEACDSTPOPCSDVA 297

```

QY 239 LANHKEVDSFRSIYTLNDGLSDSEAVAGRVYEDTNYNGNPMFLCTLLAAEQYDALYQ 358
DB 238 LANLKYVDSFRSIYTLNDGLSDSEAVAGRVYEDTNYNGNPMFLCTLLAAEQYDALYQ 357
QY 339 WDKQSLVETDVSLDFKALYSDAATGTYSSSSSTYSIVDAVKTFADGFSIVETHAAS 418
DB 338 WKAAGELANTVSLAFQOPDSSITAGTYAASSSVTSLISDIOAFDEFDIVAKYTS 417
QY 419 NSGMEQVYKSDGEQSLASDRLTWSYALLTANNRRSVVPASMGESASVGTCA--AT 476
DB 418 SFFLSEQYDKSTGADDSANLWTSYAAAITTAQAARNGFTGASWG--AAGVSTSCSTGAT 474
QY 477 SAIGTYSVTAT-----SWPSIYATG-----GTTTAT-----PTGSGSVST 514
DB 475 SFGSGSGSVETFDVYATTVYGNITITDVSBELGNMTPANGVALLSANYPTWSATIALP 534
QY 515 SKTT-----ATASKSTTTTBSG 531
DB 535 ADTTIQYKYNIDGSTVIMEDALISNREITTPASG 568

RESULT 7
Q9P4C5 PRELIMINARY; PRT; 571 AA.
AC Q9P4C5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Glucoamylase.
OS Lentinula edodes (Shiitake mushroom) (Lentinus edodes).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Tricholomataceae; Lentinula.
OX NCBI_TaxID=5353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20292826; PubMed=10831434;
RA Zhao J., Chen Y.H., Kwan H.S.;
RT "Molecular Cloning, Characterization, and Differential Expression of a
RT Glucoamylase Gene from the Basidiomycetous Fungus Lentinula edodes.";
RL Appl. Environ. Microbiol. 66:2531-2535 (2000).
DR EMBL; AF20541; AAF75523.1; -.
DR HSSP; P04064; IGAL.
DR GO; GO:0003824; F:glucan 1,4-alpha-glucosidase activity; IEA.
DR GO; GO:0004339; P:polysaccharide metabolism; IEA.
DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
DR InterPro; IPR002044; CBD_4.
DR InterPro; IPR008291; Glu-a-glycd SBD.
DR InterPro; IPR000165; Glyco_hydro_15.
DR InterPro; IPR008928; Glyco_trans_gnp.
DR Pfam; PF00686; CEM_20; 1.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR PRINTS; PR00736; GLHYDRASE15.
DR ProDom; PD001568; CBD_4; 1.
DR PIRSF; PIRSF001031; Glu-a-glycd SBD; 1.
SQ SEQUENCE 571 AA; 61161 MW; A1BD1ECC848E97FF CRC64;

Query Match 46.4%; Score 1286; DB 3; Length 571;
Best Local Similarity 47.9%; Pred. No. 6.7e-74;
Matches 266; Conservative 91; Mismatches 150; Indels 48; Gaps 10;

```

```

DB 193 TYDLMWEIYSSFFSTAAQHRALRGCTTLGRALIGQTSILASSGNGADFLCTLGYSWMT 252
QY 254 GSFTLANPDSRSGKANDATLGSITFEPDEAACDDSTOPCSPRALANHKEVDSFRSIT 313
DB 253 G-YMTANNGGSGSGIDANSVLASIHFTPDAAAGCDAITFQPCSDVALLNFTYVNAFRNAY 311
QY 314 TLNDGLSDSEAVAGRVYEDTNYNGNPMFLCTLLAAEQYDALYQMDQKSLVETDVSLD 373
DB 312 EINSGLSTNEAVLTRRYPEDVWGNPMVLTLLAAEQYDLSLWNSGSLDVTNTSLP 371
QY 374 FFKALYSDAATGTYSSSSSTYSIVDAVKTFADGFSIVETHAASNGSMSEQYDSDGEO 433
DB 372 FFKQFDSITTAGTYSSEBSFTLRAAKTFSDFVALVAKTTPSSGGLAEQISNGAP 431
QY 434 LSARDLWTSYALLTANNRRSVVPASWG-----ETSSASVPG- 471
DB 432 VSAVDLTWSYASALTAFPDARAGVIPGSGAQGLSLNCGGPVAQAVSVTFNDASTLEQ 491
QY 472 ---TCAATSAICTY-----SVTVSWPSIYATGTTT-----ATPTGSGSVST 515
DB 492 NYLITGAVDALDEWSTDNAIILSSANYPTWSVTVLPSTDVQYKRIKXDSGSGVTWESD 551
QY 516 ---KTTATASKSTT 527
DB 552 PMMETITTPANGTYAT 566

RESULT 8
Q8J0P8 PRELIMINARY; PRT; 609 AA.
AC Q8J0P8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Glucoamylase Glam.
OS Mucor circinelloides.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Mucor.
OX NCBI_TaxID=36080;
RN [1]
RP SEQUENCE FROM N.A.
RX Houghton-larsen J., Pedersen P.A.;
RL Submitted (OCT-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY168303; AAN85206.1; -.
DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
DR InterPro; IPR005036; CEM_21.
DR InterPro; IPR008928; Glyco_hydro_15.
DR InterPro; IPR000165; Glyco_trans_gnp.
DR Pfam; PF00370; CEM_21; 1.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR PRINTS; PR00736; GLHYDRASE15.
DR PROSITE; PS00820; GLUCOAMYLASE; 1.
SQ SEQUENCE 609 AA; 64778 MW; 56DD54CD371CC7C1 CRC64;

Query Match 28.3%; Score 784; DB 3; Length 609;
Best Local Similarity 39.1%; Pred. No. 8e-42;
Matches 175; Conservative 73; Mismatches 158; Indels 42; Gaps 11;

```

```

Db      342 LMEVNGVHFYTLWMVRGLVKGANFATRNQDSTRATYTTAASIKTKIDSFNNSNGQY 401
QY      257 ILANFD-----SSRGKANTL-----LGIHTFDEAACDDSTPPCSPRALANKEVYDS 308
Db      402 VVSQSVTGVGSKAGYASVLIASNLGSL-----QDCFFYPSGDKMLATVAALAESK 452
QY      309 FRSIYTLNDGLSDSEAVAVAGRYPEDTY-----YNGNPMFLCTTLAAEQLYDALYQMDKQG 363
Db      453 PASIYISINQNLNGVLAIGRYPEDTYNGNNGSQGNPFICTNPAFELYRAIKEMFNNG 512
QY      364 SLEVTVDVSLDFPKALYSDAATGT- YSSSSSTYSIYDAVKTADGFWISIVETHAASNGSM 422
Db      513 GVFTYSISLNFKKFDSAAAGTKYVGTSSFNLSLVQNVAAADAFSTIKFHAATNGSM 572
QY      423 SEQYKSDGEOLASRDLTWGAALLTAN 450
Db      573 SEQYGRADGLMTGARDLTWASHASISAS 600

```

## RESULT 9

```

0727X9 PRELIMINARY; PRT; 579 AA.

```

```

AC 0727X9; TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Glucoamylase (Fragment).
OS Rhizopus oryzae (Rhizopus delemar).
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Rhizopus.
NCBI_TaxID=64495;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen G., Yu X.C., Jiang H.Z., Li M.G.;
RL Submitted (Aug-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY359821; AAQ18643.1; -.
FT NON TER 1 26
FT CHAIN 27 85 glucoamylase.
FT CHAIN 86 411 glucoamylase.
FT CHAIN 412 579 glucoamylase.
SQ SEQUENCE 579 AA; 62151 MW; 75811063688751 CRC64;

```

Query Match 26.3%; Score 728; DB 3; Length 579;  
 Best Local Similarity 37.1%; Pred. No. 2.8e-38;  
 Matches 166; Conservative 70; Mismatches 170; Indels 42; Gaps 10;

```

QY      25 ATLDLSLNEATVARTAILNIGADGAMVSGDSGIIVASPTDNPDTYTWTRDGLV 84
Db      143 STISSWIKQEGISRFAMLRNINP-----PGSATGFIAASLSTAGPDYVAMTRDAALTS 197
QY      85 KTLVDLRR--NGDTSLSTIENTYISAQIVQGISNPSGDLSSAGAGCEPKFENVDETAYT 141
Db      198 NVIYEYENTTLISGKKTILNVLDKDYTPSVKTOSTSTVCN-----CLGEPEFNPDSGCT 251
QY      142 GSWGRPORDGALPATAMIGFGQ-WLNDNGYSTATDIPVPLVRNDLSYVQYNNQGYD 200
Db      252 GAWRPPONDGAERATTFILFADGVLQTKDASVYTGTLKRAIKRDLDYVNVVMSNGLD 311
QY      201 LMEVNGSSFFTIAVQHRALVEGSAFATAVGSS--CSWCDQADPEILCYLOSFWTGSFIL 258
Db      312 LMEVNGVHFYTLWMVRGLGLADFAKRNQDSTRASTYSTASTIANKISSFWVSNW 371
QY      259 ANFPS-----SRGKQANTL-----LGIHTFDEAACDDSTPPCSPRALANKEVYDS 308
Db      372 IQVSQSVTGVGSKAGYASVLIASNLGSL-----DDEFFTPGSEKILATVAALAESK 422
QY      309 FRSIYTLNDGLSDSEAVAVAGRYPEDTY-----YNGNPMFLCTTLAAEQLYDALYQMDKQG 363
Db      423 PASIYIPINKULPSVLGNSIGRYPEDTYNGNNGSQGNPFIFLAGYALYRAIKEMWING 482
QY      364 SLEVTVDVSLDFPKALYSDAATGT-TYSSSSSTYSIYDAVKTADGFWISIVETHAASNGSM 422

```

```

Db      483 GVTVSISLDPFFKKFDSASATSGKKYVGTSDFNNLQNLIALADRFLSTVQLAHNNGSL 542
QY      423 SEQYKSDGEOLASRDLTWGAALLTAN 450
Db      543 ABEFDRITTGARDLTWASHASISAS 570

```

## RESULT 10

```

08TFES PRELIMINARY; PRT; 515 AA.

```

```

AC 08TFES; TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Glucoamylase precursor (EC 3.2.1.3).
GN Glu 0111.
OS Saccharomycopsis fibuligera (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
NCBI_TaxID=4944;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 0111.
RA Hostinova E., Sciovicova A., Janacek S., Gasperik J.;
RT "Raw starch degrading glucoamylase from Saccharomycopsis fibuligera: molecular cloning and expression in yeast."
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ11587; CAC83969.1; -.
DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
DR InterPro; IPR00165; Glyco_hydro_15.
DR InterPro; IPR008928; Glyco_trans_6np.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR PRINTS; PR00736; GLHTRLAS15.
DR PROSITE; PS00820; GLUCOAMYLASE; 1.
KW Signal; Hydrolase; Glucosidase.
FT SIGNAL 1 26
FT CHAIN 27 515 GLUCOAMYLASE.
SQ SEQUENCE 515 AA; 57423 MW; A2F27D9EEA842AE0 CRC64;

```

Query Match 24.3%; Score 674; DB 3; Length 515;  
 Best Local Similarity 33.7%; Pred. No. 6.7e-35;  
 Matches 163; Conservative 82; Mismatches 177; Indels 62; Gaps 16;

```

QY      24 RAITDLSLNEATVARTAILNIGADGAMVSGDSGIIVASPTDNPDTYTWTRDGLV 83
Db      41 RSNFTQIHQOPAVSWYLLQNLIDYPEGQFSAKRGVAVASPTSEPDYFQWTRDAIT 100
QY      84 LKTLVDLFRN--GDTSLSTIENTYISAQIVQGISNPSGDLSS--GAGLGEPEKFNVDY 138
Db      101 FLSLIAVEVHSFENTTLAKVVEYISNTYTLQVSNPSGDFDPNHDGGEPEKFNVDY 160
QY      139 AYTSGWRPORDGALPATAMIGF-----GQMLL--DNQYT-STADIDVPLVRND 186
Db      161 AYTASWRPORDGALPATAMIGF-----GQMLL--DNQYT-STADIDVPLVRND 220
QY      187 LSYVAQVMNQGYVLMREAVNGSSFFTIAVQHRALVEGSAFATAVGSS--CSWCDQADPEI 244
Db      221 LQHVSTHMSVSGFLPMBENQTHFTALVQKALSYGIPUSKYTYNDGFTSMLEKQDAL 280
QY      245 LCYLOSFWTGSFILAND-----SSRGKQANTLGLIHTFDEAACDDSTPPCSP 296
Db      281 NSYINS---SGFVNSGKHIVESPQLSRGLSDATYIALITHD---IGDDTYTFPN- 333
QY      297 PALANKEVYDSF-----RSIYTLNDGLSDSEAVAVAGRYPEDTY-----YNGNPMFLC 344
Db      334 ---VDNSYVLSLYLLVDNMKRYKINGNY--KAGAAVAGRYPEVYVNGVGTSEGNPQOLA 388
QY      345 TLAABEQLYDALYQMDK-QGSLEVTVDVSLDFPKALYSDA--TGTYS-----SSST 393
Db      389 TAYAGQFFYTLAYNSLKNKVLTEKLYDLVNSFIADLSKIDSKDSVASKOSLTLTYGSDN 448

```



QY 334 YSSIDAVKTFADGFSIVETHAASNGMSSEQYDKSDGEQLSARDLTWSYALLTANNRR 453  
 DB 449 YKGVYKSLIQFDSFLKVLHDIDNGQLTEIINRYTGQAQAVSLTWSGSLSANRR 508  
 QY 454 NSVV 457  
 DB 509 NGLI 512

## RESULT 11

ID Q96210 PRELIMINARY; PRT; 599 AA.  
 AC Q96210;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein ST2017.  
 GN ST2017.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 NCBI\_Taxid=11955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JCM 10545 / 7;  
 RX MEDLINE=21456156; PubMed=11572479;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 Nagai Y., Nishijima K., Otsuka R., Nakamiya M., Kato Y.,  
 Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
 Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
 RA Oshima T., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic thermoacidophilic  
 RT Crenarchaeon, Sulfolobus tokodaii strain7";  
 RL DNA Res. 8:123-140(2001).  
 DR EMBL; AP000988; BAB67116.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0005976; P:polyasaccharide metabolism; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000165; Glyco\_hydro\_15.  
 DR InterPro; IPR008928; Glyco\_trans\_6hp.  
 DR InterPro; IPR000531; TonB\_boxC.  
 DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
 DR PROSITE; PS00430; TONB\_DEPENDENT\_REC\_1; 1.  
 DR Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 599 AA; 69661 MW; F8B94CME731D119 CRC64;

Query Match 6.7%; Score 187; DB 17; Length 599;  
 Best Local Similarity 22.3%; Pred. No. 0.001;  
 Matches 102; Conservative 58; Mismatches 151; Indels 146; Gaps 23;  
 QY 29 SWL-----SNEATVARTAIL-----NNIGADGAWVSGDSGIVVSPSTNDYFYWT 78  
 DB 235 AMLVAVARYDLSVRSLLIIAHMQN-----GALPALDPTDIRNNKKT-----YTVWHR 286  
 QY 79 DSGVLKTLVDFRNGDTSLSTIENYISAQIVOGISNPSGDLSSGAGLGPKNVDET 138  
 DB 287 DAAF-----ASIALTLGY-----QDPINLNFMTKPLFNGFLPQ-KYTC-- 327  
 QY 139 AYTGWKRP-----QRDGPALRAITAMIGCGWMLDNGYSTNTDIYWPLVRDL 187  
 DB 328 ---GNWGSTWHPWNRSLPIOD-----ETALMLALWVHFSRFTD--IDFVRPLYPABV 377  
 QY 188 SYVAQY-----NOTG-----YDLWEVNGSSFTTIAVOHRAIVGSAFATAVGSSGMC 237  
 DB 378 KKIAEFLVSYNDEEGPLPSYDLMEERLGTHTFTSLAVYAGLMAKYFAEFGBE-NLK 436  
 QY 238 D---SQAPEILCYLOSFWTGS-FILANFDSRSRGK--DANTLLGSIHTFPDEACDDSTF 291  
 DB 437 DKYLTAAEIVKKGIERFVGVGHFAITFYIDENSIDKTVASTLFAAIL-----GPF 486

QY 292 QPCSPRALANKEVVDSPRSITYLNDGLDSEAVAGRYPEPTY-----NGNPFICTLA 347  
 DB 487 DPKDPRVSNRKYVEBK-----NING-----IAIYENDWYLKODEKSANAFITLIM 534  
 QY 348 AAEQLYDALYOWDKGSLVTVDSLDFPKALYSDAATGYSSSSSTYSIYDAVKTADG 407  
 DB 535 LAQY-----ILEGNKKA 550  
 QY 408 FVSIYETPAASNGMSSEQYDKSDGEQLSARDLTWSYA 444  
 DB 551 YIDWVISHMLPTGILPEQVSPKN-TYPSVAPLWMSHA 586

## RESULT 12

ID Q9HL88 PRELIMINARY; PRT; 636 AA.  
 AC Q9HL88;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein Ta0342.  
 GN Ta0342.  
 OS Thermoplasma acidophilum.  
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
 OC Thermoplasmataceae; Thermoplasma.  
 NCBI\_Taxid=2303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 1728;  
 RX MEDLINE=20479972; PubMed=11029001;  
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,  
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;  
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma  
 RT acidophilum";  
 RL Nature 407:508-513(2000).  
 DR EMBL; AL445064; CAC11486.1; -;  
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.  
 DR GO; GO:0005976; P:polyasaccharide metabolism; IEA.  
 DR InterPro; IPR000165; Glyco\_hydro\_15.  
 DR InterPro; IPR008928; Glyco\_trans\_6hp.  
 DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
 DR Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 636 AA; 73269 MW; FFA4656CCDB3A9FD CRC64;

Query Match 6.7%; Score 185.5; DB 17; Length 636;  
 Best Local Similarity 19.0%; Pred. No. 0.0014;  
 Matches 86; Conservative 56; Mismatches 139; Indels 171; Gaps 15;  
 QY 57 DSGIVVSPSTND-----NPDYFYWTFRDSGLVLTVDLFRNGDTS--LSTIENYISAQ 109  
 DB 281 DLGALNASCDDLDLKLSDGYGYVWPBPDSMAAYAL-SIGHSSTARRFFALMEDSLSE 339  
 QY 110 AIVOGISNPSGDLSSGAGLGPKNVDETAYTGSWGRPQDGPAL-----BATAWIGGQ 164  
 DB 340 GYLHAKNVDEKIAS-----SWLPHVWNGKSLYPIODEBTALVVAL 381  
 QY 165 WLDNGYSTATDIWPLVRDLSYVAQYN-----QTGYDLWEE 204  
 DB 382 WEYFRKY-----NDIGFTAPYRRLITRADPMTNFVNDNGLPKPSFDLWEE 428  
 QY 205 VNGSSFTTIAVOHRAIVGSAFATAVGSSGMCWSQAPILICYLOSFWTGSFILANFSS 264  
 DB 429 RYGIHAYVAVVYALKAASNFANVG-----DPD 458  
 QY 265 RSGKDANTLLGSIHTFPEAACDDSTFQPCSPRALANKE--VVDSPRSITYLNDGLSDS 322  
 DB 459 LSEKEMNAEEMNYAFDRFSEDTGY--ARALIIOCKPFTYDLSALTSVL-FGMDA 514  
 QY 323 E-----AVAVGRYPEDTYN-----GNPFLCTLAABEQLYDA 355  
 DB 515 DDPKIVSTORISEDLMWNGVGIAFYONDRYKRVKODPSVPGNPMITTLMAA-RYYMR 573

```

QY 356 LYQMDKQGSLEVTIVSLDFFKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETH 415
DB 574 FGDEKAMNL-----IOWKSH 590
QY 416 AASGMSSEQYDKSDQGLSARDLTWSYALL 447
DB 591 RQKSGIFSEQINPYNGEPLSVPLWSHSEPI 622

RESULT 13
08XDG4
ID 08XDG4 PRELIMINARY; PRT; 1588 AA.
AC Q8XDG4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative adhesin (Hypothetical protein).
GN Z5029 OR EC54480 OR H161.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. II, Burland V., Mau B., Glaesner J.D.,
RA Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Poefel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Alathararaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsundo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 0509952;
RA Makino K., Yutsundo C.H., Yokoyama K., Kubota Y., Kimura S.,
RA Shinagawa H.;
RT "O157 specific gene similar to H. influenzae adhesin gene.";
RL Submitted (DSC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF002556; BAB37903.1; -
DR EMBL; AB036416; BAB87814.1; -
DR PIR; A86036; A86036.
DR PIR; H91188; H91188.
DR InterPro; IPR008640; Hep_Hag.
DR InterPro; IPR008635; HIM.
DR InterPro; IPR005594; Yada.
DR Pfam; PF05658; Hep_Hag; 13.
DR Pfam; PF03895; HIM; 12.
DR Pfam; PF03895; Yada; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 1588 AA; 160150 MM; B2BA4B06EFPF28DEC CRC64;

Query Match 6.5%; Score 181; DB 16; Length 1588;
Best Local Similarity 21.7%; Pred.No. 0.0092;
Matches 136; Conservative 95; Mismatches 217; Indels 178; Gaps 30;
QY 15 TGLANVSKRTLDSSWLSNEATVARTALNNIGADG-AWVGAGSGIVAVSPSTNDNDFY 73
DB 576 TTNANNTSNATNTNTNISN-----LTETVTNLGEBALKM--DKONGVFTAAHGTEETSKI 628

```

```

QY 74 YTWTRD-----SGLVLTIVDLFRNGDTSLSITENYISAQAIVOGISNPSGDL 122
DB 629 -TNVKGODLTGSGTDAVANGSLKTTNDVAVTNTNTNATNTN-----ISNLTETV 677
QY 123 SSGAGLGEF--KRVNDETATYGSWGRPQRDGPALRATAMFGQMLDNGTSTATIWM 180
DB 678 TN--LGEDALKMDKMGVFTAAHG--NNTASKITN-----ILDGVVTATSSDAIN 723
QY 181 PLVRNDL-SYVAQW-----NQTG-----YDMEEVNGSFFTIIVGRALVEGSAFA 227
DB 724 GSQLYDLSNIAIATYFGNNAVNTDGVPTGPTYIGE-----TNTYVGDALAI--NSFS 777
QY 228 TAVGSSCSWCDSQLPILCYLQSFMTGSFI--LANPDSRSKGA--NTLLGSIHF-- 280
DB 778 TSLGDALWM-DATAGKPSAHGNTGDSVITVDAGDISDSSDAVANGSQLHGVSSVVD 836
QY 281 ---DPEAACDDSTFQCSFRALANKEVDSFRST-YTLNDGLSDSEAVANGRIPEPTY 335
DB 837 ALGGGAENVNADGTITTAPTYTIANADYDNDGALNALIDTTLDDAL-----LMDADAG 887
QY 336 YNGNPWFICTLAAAEQLYDALYQMDKQGSLEVTIVSLDFFKALYSDAATG----- 385
DB 888 ENG-----AFSAHGKDKTASV--TNVANGAISASDSALINSQLYTTNKYI 933
QY 386 -----TYSSSSTYSIVDAVKTF-----A 405
DB 934 ADALGGDAENVADGTITAPPTTIANAEYNVGALDLDNALLMDETANGAGAYNASH 993
QY 406 DGFVSIYETHAANGSNSE-QYDKSDGEQLSA-----RDLTWSYALL 446
DB 994 DGRASITTN--VANGSISEDSTDAVNGSQLNATNMMEONTQIINOLAGNTDATTYIGENG 1051
QY 447 LTANNRNSVVPASWGEITSASVPGTCATSAIGTSSVTVTSPISVATGCTTTATPT 506
DB 1052 AGINVTATNDGALFANDASHQVGAT-----ALG-YNSVAKGSSVAIGGYSVDVTGI 1105
QY 507 GSGSVTSTSKTTTATASKTSITTRSGM 532
DB 1106 ALGSSSVSSVIAKSGSDTSTENG 1131

RESULT 14
09C105
AC 09C105 PRELIMINARY; PRT; 1236 AA.
ID 09C105;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative glucosylase I (Alpha-1,4-glucan glucosidase), extracellular
DE starch-degrading enzyme, by similarity to S. cerevisiae STAl, contains
DE chitinase family signature.
GN SPABIE7.04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Wood V., Rajandream M.A., Barrell B.G., Seeger K., Harris D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL590605; CAC36921.1; -
DR GenesD_Sbombe; SPABIE7.04C; -
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001917; Aminoctans II.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
SQ SEQUENCE 1236 AA; 123387 MM; 5A2D33A30B87CDD8 CRC64;

```

Query Match 6.3%; Score 175; DB 3; Length 1236;  
Best Local Similarity 21.5%; Pred. No. 0.016;  
Matches 129; Conservative 91; Mismatches 242; Indels 136; Gaps 21;

```
QY 10 SGLVCTGLANVISKRAITLDSWLSNEATVARTALINNIGAD----- 49
D 528 SSIISPMTSVLSSSSIPTSSSDPSSSITTISSGSSSSIPSTFSSVSSLSSTSP 587
QY 50 ---GAWVGADBSGIYVAPSTNDPFIYTWTDGLVLTVDLFRNGDTLSLTENYI 106
D 568 SSTLSISSTSTSTFSSASTSPSSISSSSSTLSSPP-----STSSIMI 637
QY 107 SQAIVVIGISN-----PSGD--LSSGAGLGEPEKENVETAYT 141
D 638 SSSSISSSSSLSLSSSITPTISSSLSTSSSVTPSSSTLVSSSSL-----IVSSPVA 692
QY 142 GSWGRPDGPRALATAMIGFGQWLLDNGYTSATDIWPLVRNDLSYVAQYMNQGYDL 201
D 693 SSSSPFIPSSSLVST-----YSASLSNI-----THSSLSLTAM--SSSAI 732
QY 202 MEVVGSSFFTIYVGRHALVEGSAFATVAGSS--CSWCDQAPPELLCYLQSFWTGSFLIA 259
D 733 PTSVNSSLTIT-ASSSNTLLSSITSSSAIVSSTTVSNISNLPATASQSLTNSSTLA 791
QY 260 N--FDSRSRGKDNLTLSIHTEDEACDDSTFQPCSPRALANKEVVDSPRSIYTLND 317
D 792 TSLVSSSSSRITISSTINEVTSFHALPTVSTTSSSTSLANKGV--NSNSTSLN- 848
QY 318 GLDSEAVAVGRYPEDTYNGNPFCLTLAAEQLYDALYQMDKQSLVTDVSLDFPKA 377
D 849 -LESTSVSTAYTTDSVTS-----TTALTSQ-----GPSSSVSSLSSTTS 890
QY 378 L-----YSDATGYSSSSSTYSIVD-----AVKTFADGFVSIYET 414
D 891 LSTSIPTVTSVAPATVTSGETSTSVYGSCTDSATSSWTAEISSAITSVAASYTPRSS 950
QY 415 HAASNGSEQDYKDGQSLARDLTWGYAALLTANNRNSVPAKMETASASVPG--- 471
D 951 SSSASWSSSEVDSTASATAGSS-TSISATASVGSSTSVATASATDSIAAASV 1009
QY 472 TCAATSAIGTYSVTVTSWPSIVATGCTTTATPTGSGSVT-STSKTATASKTSTTTS 530
D 1010 TGSSSTSVAT-ASVTDSSSTSVATASATDSSTSSIAVASVYGSSTSVATASATDSISS 1066
```

RESULT 15

```
Q97BM7 PRELIMINARY; PRT; 659 AA.
AC 097BM7;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Glucoamyase.
GN TV0428 OR TVG0415361.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasma; Thermoplasmales;
OC Thermoplasmataceae; Thermoplasma.
CX NCBI_TaxID=50339;
PP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS1 / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=1121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Nunobiki T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.,
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanium."
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
DR EMBL, AF000992; BAB59570.1;
DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
DR InterPro; IPR000165; Glyco_hydro_15.
```

DR InterPro; IPR008928; Glyco\_trans\_6hp.  
DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
KW Complete proteome.  
SQ SEQUENCE 659 AA; 76598 MW; 36101F5405F13B33 CRC64;

Query Match 6.3%; Score 174.5; DB 17; Length 659;  
Best Local Similarity 19.7%; Pred. No. 0.0071;  
Matches 90; Conservative 49; Mismatches 150; Indels 159; Gaps 16;

```
QY 57 DSGIVVAPSTND-----NPDIYTWTRDSGLVLTVDLFRNGDTLSLTENYISAQAI 111
D 304 DLGAIAACDDLDLMSHDGYVWPRDASIAAYAL-----SIGHSETARFPAL 354
QY 112 VQGISNPSGDLSSGAGLGEPEKENVETAYTSW-----GRP-----QRDGPRLATAMIG 161
D 355 MESSKSEEGYLH-----KKNVD-GKIASSWLPHVINGKHLYIQDEDTAL----- 399
QY 162 FGQWLLDNGYTSATDIWPLVR-----NDLSYVAQYMN-----Q 196
D 400 -----VWVLMSEYFHKNDIGFTAPYEKLTAKADEMTFRVDSGLPQ 443
QY 197 TGYDMEVVGSSFFTIYVGRHALVEGSAFATVAGSSCSWCDQAPPELLCYLQSFWTGSF 256
D 444 PSFDLMEERYGVHAYTVATVYALKAASNFAVFGDG-----DLSEKYLISAERYEA-- 496
QY 257 ILANFDSRSGKDNLTLSIHTEDEACDDSTF-----QPCSPRALANKEVVDSPRSI 312
D 497 FEEKYSESEYVAPATIDGKPDFTVDALTSVIFGKMDPRDPKIVSTMEKI----- 549
QY 313 YTLNDGLSDSEAVAVGRYPEDTYN-----GNPFLCTLAAG--QYDALYQMDK 361
D 550 ---SDTLMVNNGVGARQNRKVRKVDQTNIPGNPAILITLTMARYLLRGDPERAMD- 605
QY 362 QGSLVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTADGFVSIYETHAASNGS 421
D 606 -----LIMVYKSHRQKGI 619
```

Search completed: June 17, 2004, 17:47:29  
Job time : 50 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:36:33 ; Search time 18 Seconds

(without alignments)  
1544.748 Million cell updates/sec

Title: US-10-038-723-2

Perfect score: 2771

Sequence: 1 MSFSLALSLGLVCTGLANV.....SKTTATASKTSTTTSGMSL 534

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2742	99.0	640	1	AMYG ASPNG
2	2620.5	94.6	639	1	AMYG ASPSH
3	2588.5	93.4	639	1	AMYG ASPKA
4	1935	69.8	612	1	AMYG ASPOR
5	1577	56.9	626	1	AMYG NEUCR
6	1321	47.7	616	1	AMYG HORRE
7	781.5	26.2	450	1	AMYG SCHPO
8	725	26.2	604	1	AMYG RHIOF
9	671.5	24.2	519	1	AMYG SACFI
10	664.5	24.0	519	1	AMYG SACFI
11	641.5	22.2	624	1	AMYG ARXAD
12	597	21.5	549	1	AMYG YEAST
13	522	18.8	767	1	AMTH SACDI
14	521	18.8	768	1	AMTH SACDI
15	221.5	8.0	615	1	YGI0 METUA
16	159.5	5.8	1196	1	ICGV_PSEEX
17	159	5.7	1210	1	ICEN_PSEEL
18	158	5.7	670	1	YFEG SCHPO
19	158	5.7	3178	1	YF89 CAEBL
20	157	5.7	1258	1	ICEN ERHME
21	156.5	5.6	540	1	GXU1 ASPAC
22	155.5	5.6	1148	1	ICER_PSEEX
23	155	5.6	1122	1	ICBA PANAN
24	154.5	5.6	1200	1	ICEN_PSEEX
25	151.5	5.5	1034	1	ICEN PANAN
26	151.5	5.5	1306	1	MSB2 YEAST
27	151.5	5.5	1331	1	MANB CALSA
28	150	5.4	1609	1	FIQ2 YEAST
29	147	5.3	507	1	Y646 YEAST
30	142	5.1	1419	1	ALAI CANAL
31	141.5	5.1	644	1	XYND CELFI
32	141	5.1	1802	1	HKRI YEAST
33	140.5	5.1	1567	1	ICEN_XANCT

34	138.5	5.0	600	1	SP96 DICDI
35	138	5.0	797	1	VGIX HSVB
36	137	4.9	556	1	MSG3 YEAST
37	136	4.9	542	1	CH12 RHIOF
38	136	4.9	995	1	Y109 YEAST
39	135.5	4.9	537	1	GU1 PENJA
40	133.5	4.8	1589	1	PHP DROME
41	132	4.8	827	1	XANP XAN62
42	130.5	4.7	1537	1	FIQ1 YEAST
43	130.5	4.7	1746	1	TENA FIG
44	130	4.7	706	1	PLB2 YEAST
45	129.5	4.7	1367	1	AMTH YEAST

## ALIGNMENTS

RESULT 1

ID	AMYG ASPNG	STANDARD	PRT	640 AA.
AC	P04064; Q92201; Q99179;			
DT	01-NOV-1986 (Rel. 03, Created)			
DT	01-NOV-1986 (Rel. 03, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Glucosylase G1 and G2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucosylhydrolase).			
GN	GLA.			
OS	Aspergillus niger, and			
OC	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=5061, 105351;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=A.niger;			
RX	MEDLINE=84261458; PubMed=6204865;			
RA	Boel E., Hansen M.T., Hjort I., Hoegh I., Pii N.P.;			
RT	"Two different types of intervening sequences in the glucosylase gene from Aspergillus niger."			
RL	EMBO J. 3:1581-1585(1984).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=A.niger;			
RX	MEDLINE=84236105; PubMed=6203744;			
RA	Boel E., Hjort I., Svensson B., Norris F., Norris K.E., Pii N.P.;			
RT	"Glucosylase G1 and G2 from Aspergillus niger are synthesized from two different but closely related mRNAs."			
RL	EMBO J. 3:1097-1102(1984).			
RN	[3]			
RP	SEQUENCE OF 25-640, AND COMPARISON OF FORMS G1 AND G2.			
RC	SPECIES=A.niger;			
RX	MEDLINE=86136085; PubMed=3081341;			
RA	Svensson B., Larsen K., Gunnarsson A.;			
RT	"Characterization of a glucosylase G2 from Aspergillus niger."			
RL	Bur. J. Biochem. 154:497-502(1986).			
RN	[4]			
RP	SEQUENCE OF 25-640.			
RC	SPECIES=A.niger;			
RA	Svensson B., Larsen K., Svendsen I., Boel E.;			
RT	"The complete amino acid sequence of the glycoprotein, glucosylase G1, from Aspergillus niger."			
RL	Carlsberg Res. Commun. 48:529-544(1983).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=A.awamori;			
RX	MEDLINE=85085934; PubMed=6440004;			
RA	Numberg J.H., Meade J.H., Cole G., Lawyer F.C., McCabe P.;			
RT	Schweikart V., Tal R., Wittman V.P., Flataard U.E., Innis M.A.;			
RL	"Molecular cloning and characterization of the glucosylase gene of Aspergillus awamori."			
RN	Mol. Cell. Biol. 4:2306-2315(1984).			
RP	[6]			
RP	REVISIONS.			

RC SPECIES=A.awamori;  
 RA Nunberg J.H., Meade J.H., Cole G., Lawler F.C., McCabe P.,  
 RA Schweickart V., Tal R., Wiltman V.P., Flatgaard J.B., Innis M.A.;  
 RL Submitted (FEB-1985) to the EMBL/Genbank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 1-11 FROM N.A.  
 RC SPECIES=A.niger; STRAIN=ATCC 10864;  
 RX MEDLINE=91168302; PubMed=2076554;  
 RA Fowler T., Berka R.M., Ward M.;  
 RT "Regulation of the gla gene of *Aspergillus niger*.";  
 RL Curr. Genet. 18:537-545(1990).  
 RN [8]  
 RP CONFORMATION OF O-GLYCOSYLATED REGION.  
 RX MEDLINE=92189576; PubMed=1546955;  
 RA Williamson G., Belshaw N.J., Williamson M.P.;  
 RT "O-glycosylation in *Aspergillus* glucanase. Conformation and role  
 RT in binding."; Biochem. J. 282:423-428(1992).  
 RL [9]  
 RN ACTIVE SITES, AND MUTAGENESIS.  
 RC SPECIES=A.awamori;  
 RX MEDLINE=90231978; PubMed=1970434;  
 RA Sierks M.R., Ford C., Reilly P.J., Svensson B.;  
 RT "Catalytic mechanism of fungal glucanase as defined by mutagenesis  
 RT of Asp176, Glu179 and Glu180 in the enzyme from *Aspergillus*  
 RT awamori."; Protein Eng. 3:193-198(1990).  
 RL [10]  
 RN MUTAGENESIS OF TRP-144.  
 RC SPECIES=A.awamori;  
 RX MEDLINE=90046622; PubMed=2510150;  
 RA Sierks M.R., Ford C., Reilly P.J., Svensson B.;  
 RT "Site-directed mutagenesis at the active site Trp120 of *Aspergillus*  
 RT awamori glucanase."; Protein Eng. 2:621-625(1989).  
 RL [11]  
 RN MUTAGENESIS.  
 RC SPECIES=A.awamori;  
 RX MEDLINE=93165653; PubMed=8433972;  
 RA Sierks M.R., Ford C., Reilly P.J., Svensson B.;  
 RT "Functional roles and substrate locations of Leu177, Trp178 and Asn182  
 RT of *Aspergillus* awamori glucanase determined by site-directed  
 RT mutagenesis."; Protein Eng. 6:75-79(1993).  
 RL [12]  
 RN CHARACTERIZATION OF CATALYTIC DOMAIN.  
 RC SPECIES=A.awamori;  
 RX MEDLINE=93277459; PubMed=8503847;  
 RA Stoffer B., Frandsen T.P., Busk P.K., Schneider P., Svendsen I.,  
 RA Svensson B.;  
 RT "Production, purification and characterization of the catalytic  
 RT domain of glucanase from *Aspergillus niger*."; Biochem. J. 292:197-202(1993).  
 RL [13]  
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 25-495.  
 RC SPECIES=A.awamori; STRAIN=Var. X100;  
 RX MEDLINE=92406872; PubMed=1527049;  
 RA Aleshin A., Golubev A., Firssov L.M., Honzatko R.B.;  
 RT "Crystal structure of glucanase from *Aspergillus* awamori var. X100  
 RT to 2.2-A resolution."; J. Biol. Chem. 267:19291-19298(1992).  
 RL [14]  
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 25-495.  
 RC SPECIES=A.awamori; STRAIN=Var. X100;  
 RX MEDLINE=94253149; PubMed=8195212;  
 RA Aleshin A., Firssov L.M., Honzatko R.B.;  
 RT "Refined structure for the complex of acarbose with glucanase from  
 RT *Aspergillus* awamori var. X100 to 2.4-A resolution."; J. Biol. Chem. 269:15631-15639(1994).  
 RL [15]  
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 25-495.  
 RC SPECIES=A.awamori; STRAIN=Var. X100;  
 RX MEDLINE=94231577; PubMed=8176747;  
 RA Aleshin A., Hoffman C., Firssov L.M., Honzatko R.B.;

RT "Refined crystal structures of glucanase from *Aspergillus* awamori  
 RT var. X100."; J. Mol. Biol. 238:575-591(1994).  
 RL [16]  
 RN STRUCTURE BY NMR OF 533-640.  
 RX MEDLINE=96266494; PubMed=8683599;  
 RA Sorimachi K., Jacke A.J., Le Gal-Coeffet M.-F., Williamson G.,  
 RA Archer D.B., Williamson M.P.;  
 RT "Solution structure of the granular starch binding domain of  
 RT glucanase from *Aspergillus niger* by nuclear magnetic resonance  
 RT spectroscopy."; J. Mol. Biol. 259:970-987(1996).  
 RL [17]  
 RN STRUCTURE BY NMR OF 533-640.  
 RX MEDLINE=97341228; PubMed=9195884;  
 RA Sorimachi K., Le Gal-Coeffet M.-F., Williamson G., Archer D.B.,  
 RA Williamson M.P.;  
 RT "Solution structure of the granular starch binding domain of  
 RT *Aspergillus niger* glucanase bound to beta-cyclodextrin."; Structure 5:647-661(1997).  
 RL [18]  
 RN CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-  
 CC glucose residues successively from non-reducing ends of the chains  
 CC with release of beta-D-glucose.  
 CC [19]  
 CC ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=G1;  
 CC IsoId=P04064-1; Sequence=Displayed;  
 CC Name=G2;  
 CC IsoId=P04064-2; Sequence=VSP\_000262;  
 CC [20]  
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.  
 CC [21]  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC [22]  
 CC EMBL; X00548; CAA25219.1; -  
 CC EMBL; X00712; CAA25303.1; -  
 CC EMBL; X00712; CAA25304.1; -  
 CC EMBL; K02465; AAB59296.1; -  
 CC EMBL; K02465; AAB59297.1; -  
 CC EMBL; X56442; CAA39825.1; -  
 CC PIR; A29166; A29166.  
 CC PIR; A90986; A90986.  
 CC PDB; 1AGW; 30-SEP-94.  
 CC PDB; 1GLW; 31-JUL-94.  
 CC PDB; 3GLY; 01-NOV-94.  
 CC PDB; 1DOG; 30-APR-94.  
 CC PDB; 1AC0; 07-JUL-97.  
 CC PDB; 1ACZ; 07-JUL-97.  
 CC PDB; 1GAH; 17-AUG-96.  
 CC PDB; 1GAI; 17-AUG-96.  
 CC PDB; 1KUL; 11-JUL-96.  
 CC PDB; 1KUM; 11-JUL-96.  
 CC Glycosylated; P04064; -  
 CC InterPro; IPR002044; CBD\_4.  
 CC InterPro; IPR008291; Glu-a-glcSD\_SBD.  
 CC InterPro; IPR000165; Glyco\_hydro\_15.  
 CC InterPro; IPR008928; Glyco\_trans\_6hp.  
 CC Pfam; PF00686; CBM\_20; 1.  
 CC Pfam; PF00723; Glyco\_hydro\_15; 1.  
 CC PRINTS; PRS001031; Glu-a-glcSD\_SBD; 1.  
 CC PRINTS; PRS00736; GLYDRLASE15.  
 CC PRODOM; PD001568; CBD\_4; 1.  
 CC PROSITE; PS00820; GLUCAMYLASE; 1.  
 CC Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;  
 KW Alternative splicing; Signal; 3D-structure.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT PROPEP 19 24  
 FT CHAIN 25 640 GLUCAMYLASE G1.

Query Match 99.0%; Score 2742; DB 1; Length 640;  
 Best Local Similarity 99.6%; Pred. No. 9.8e-174;  
 Matched 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSFRLSLALSGVCTGLANVSKRATLDSMLSEATVARTALINNIGDGVAVSGADSGI 60  
 DB 1 MSFRLSLALSGVCTGLANVSKRATLDSMLSEATVARTALINNIGDGVAVSGADSGI 60  
 QY 61 VVASPTNDPDPFYWTTRDSGLVLTVDLFRNGDTSLSLTENYISQAQIVQGISNPSG 120  
 DB 61 VVASPTNDPDPFYWTTRDSGLVLTVDLFRNGDTSLSLTENYISQAQIVQGISNPSG 120  
 QY 121 DLSGAGLGEPEKFNDETAATYTSWGRPQDGPALRATAMIGFQWLDNGYSTATDIW 180  
 DB 121 DLSGAGLGEPEKFNDETAATYTSWGRPQDGPALRATAMIGFQWLDNGYSTATDIW 180  
 QY 181 PLVRNDLSVVAQYNNQTYDLMEEVNGSFFTIAYQHRALVEGSAFATAVSGSSCWCSQ 240  
 DB 181 PLVRNDLSVVAQYNNQTYDLMEEVNGSFFTIAYQHRALVEGSAFATAVSGSSCWCSQ 240  
 QY 241 APEILCYLOSFWTGSFILANFDSRSGKANTLGSHTFDPBAACDSTFQPCSPRALA 300  
 DB 241 APEILCYLOSFWTGSFILANFDSRSGKANTLGSHTFDPBAACDSTFQPCSPRALA 300  
 QY 301 NHKEVVDSPRSITYLTNDGLDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360  
 DB 301 NHKEVVDSPRSITYLTNDGLDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360  
 QY 361 KQGSLEVTDLDFPKALYSDAATGYSSSSSTYSSIVAVTFPADGFSIVETHAASNG 420  
 DB 361 KQGSLEVTDLDFPKALYSDAATGYSSSSSTYSSIVAVTFPADGFSIVETHAASNG 420  
 QY 421 SMSEQYDSDGEQLSARDLTMSYALLTANNRRNSVVPASWETSASVPTCAATSAIG 480  
 DB 421 SMSEQYDSDGEQLSARDLTMSYALLTANNRRNSVVPASWETSASVPTCAATSAIG 480  
 QY 481 TYSSVTWSPSIYVATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530  
 DB 481 TYSSVTWSPSIYVATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530

RESULT 2  
 AMYG\_ASPIH STANDARD; PRT; 639 AA.  
 ID AMYG\_ASPIH  
 AC P22832;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Glucanase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)  
 DE (1,4-alpha-D-glucan glucohydrolase).  
 GN GLA.  
 OS Aspergillus shirousami.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiiales; Trichocommataceae; mitosporic Trichocommataceae; Aspergillius.  
 OC NCBI\_TaxID=5070;  
 RX MEDLINE=91182400; PubMed=1368603;  
 RA Shibusawa I., Gomi K., Iimura Y., Takahashi K., Tamura G., Hara S.;  
 RT "Molecular cloning of the glucanase gene of Aspergillus shirousami  
 and its expression in Aspergillus oryzae.";  
 RL Agric. Biol. Chem. 54:1905-1914(1990).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-  
 glucose residues successively from non-reducing ends of the chains  
 with release of beta-D-glucose.  
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC +-----+  
 CC EMBL: D10460; BAA01254.1; -  
 CC DR HSSP; P04064; IGA1.  
 CC DR InterPro; IPR002044; CBD 4.  
 CC DR InterPro; IPR008291; Glu-a-glucd\_SBD.  
 CC DR InterPro; IPR001651; Glyco\_hydro\_15.  
 CC DR InterPro; IPR008928; Glyco\_trans\_4hp.  
 CC DR Pfam; PF00686; CEM\_20; 1.  
 CC DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
 CC DR PRINTS; PIRSF001031; Glu-a-glucd\_SBD; 1.  
 CC DR PRODOM; PDOM01568; CBD 4; 1.  
 CC DR PROSITE; PS00820; GLUCOMYLASE; 1.  
 CC KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;  
 CC Signal.  
 CC FT SIGNAL; 1 18 POTENTIAL.  
 CC FT PROPEP; 19 24 BY SIMILARITY.  
 CC FT CHAIN; 25 639 GLUCOMYLASE.  
 CC FT BINDING; 143 143 SUBSTRATE (BY SIMILARITY).  
 CC FT ACT\_SITE; 199 199 CATALYTIC BASE (BY SIMILARITY).  
 CC FT ACT\_SITE; 202 202 GENERAL ACID CATALYST (BY SIMILARITY).  
 CC FT ACT\_SITE; 203 203 INTERACT WITH SUBSTRATES (BY SIMILARITY).  
 CC FT DISTULFD; 233 236 BY SIMILARITY.  
 CC FT DISTULFD; 245 472 BY SIMILARITY.  
 CC FT DISULFD; 285 293 BY SIMILARITY.  
 CC FT CARBOHYD; 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD; 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD; 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SQ SEQUENCE 639 AA; 68130 MW; E93DAE55ED72326 CRC64;

Query Match 94.6%; Score 2620.5; DB 1; Length 639;  
 Best Local Similarity 94.3%; Pred. No. 1e-165;  
 Matches 500; Conservative 16; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSFRLSLALSGVCTGLANVSKRATLDSMLSEATVARTALINNIGDGVAVSGADSGI 60  
 DB 1 MSFRLSLALSGVCTGLANVSKRATLDSMLSEATVARTALINNIGDGVAVSGADSGI 60  
 QY 61 VVASPTNDPDPFYWTTRDSGLVLTVDLFRNGDTSLSLTENYISQAQIVQGISNPSG 120  
 DB 61 VVASPTNDPDPFYWTTRDSGLVLTVDLFRNGDTSLSLTENYISQAQIVQGISNPSG 120  
 QY 121 DLSGAGLGEPEKFNDETAATYTSWGRPQDGPALRATAMIGFQWLDNGYSTATDIW 180  
 DB 121 DLSGAGLGEPEKFNDETAATYTSWGRPQDGPALRATAMIGFQWLDNGYSTATDIW 180  
 QY 181 PLVRNDLSVVAQYNNQTYDLMEEVNGSFFTIAYQHRALVEGSAFATAVSGSSCWCSQ 240  
 DB 181 PLVRNDLSVVAQYNNQTYDLMEEVNGSFFTIAYQHRALVEGSAFATAVSGSSCWCSQ 240  
 QY 241 APEILCYLOSFWTGSFILANFDSRSGKANTLGSHTFDPBAACDSTFQPCSPRALA 300  
 DB 241 APEILCYLOSFWTGSFILANFDSRSGKANTLGSHTFDPBAACDSTFQPCSPRALA 300  
 QY 301 NHKEVVDSPRSITYLTNDGLDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360  
 DB 301 NHKEVVDSPRSITYLTNDGLDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360  
 QY 361 KQGSLEVTDLDFPKALYSDAATGYSSSSSTYSSIVAVTFPADGFSIVETHAASNG 420  
 DB 361 KQGSLEVTDLDFPKALYSDAATGYSSSSSTYSSIVAVTFPADGFSIVETHAASNG 420  
 QY 421 SMSEQYDSDGEQLSARDLTMSYALLTANNRRNSVVPASWETSASVPTCAATSAIG 480  
 DB 421 SMSEQYDSDGEQLSARDLTMSYALLTANNRRNSVVPASWETSASVPTCAATSAIG 480  
 QY 480 TYSSVTWSPSIYVATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 529  
 DB 480 TYSSVTWSPSIYVATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 529





RESULT 4  
AMYG ASPOR STANDARD; PRT; 612 AA.  
AC P36914;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Glucosylase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)  
DE (1,4-alpha-D-glucan glucanohydrolase).  
GN GLA.  
OS Aspergillus oryzae.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eucristales; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.  
OX NCBI\_TaxID=5062;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91254744; PubMed=1368680;  
RA Hara Y., Kitamoto K., Gomi K., Kumagai C., Tamura G., Hara S.;  
RT "The glucosylase cDNA from Aspergillus oryzae: its cloning,  
RT nucleotide sequence and expression in Saccharomyces cerevisiae";  
RL Agric. Biol. Chem. 55:941-949(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92104497; PubMed=1761224;  
RA Hara Y., Tsuchiya K., Kitamoto K., Gomi K., Kumagai C., Tamura G.,  
RT "Nucleotide sequence and expression of the glucosylase-encoding gene  
RT (gluc) from Aspergillus oryzae.";  
RL Gene 108:145-150(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=RIB 40;  
RA Hara S., Tamura G., Kumagai C., Gomi K., Kitamoto K., Tsuchiya K.,  
RA Hara Y.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-  
CC glucose residues successively from non-reducing ends of the chains  
CC with release of beta-D-glucose.  
CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D01035; BA00941.1; -  
DR EMBL; D10698; BA01540.1; -  
DR PIR; J01346; J01346.  
DR HSSP; P04064; IGA1.  
DR InterPro; IPR002044; CBD 4.  
DR InterPro; IPR008291; Glu-a-glucd\_SBD.  
DR InterPro; IPR000165; Glyco\_hydro\_15.  
DR InterPro; IPR008928; Glyco\_trans\_6hp.  
DR Pfam; PF00686; CBM\_20; 1.  
DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
DR PIRSF; PIRSF001031; Glu-a-glucd\_SBD; 1.  
DR PRINTS; PR00736; GLHYDRLASE15.  
DR ProDom; PD001568; CBD 4; 1.  
DR PROSITE; PS00820; GLUCOMYLASE\_1.  
KM Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;  
KM Signal.  
FT SIGNAL 1 19 BY SIMILARITY.  
FT PROPEP 20 25 BY SIMILARITY.  
FT CHAIN 26 612 GLUCOMYLASE.  
FT BINDING 146 146 SUBSTRATE (BY SIMILARITY).  
FT ACT\_SITE 202 202 CATALYTIC BASE (BY SIMILARITY).  
FT ACT\_SITE 205 205 GENERAL ACID CATALYST (BY SIMILARITY).  
FT ACT\_SITE 206 206 INTERACT WITH SUBSTRATES (BY SIMILARITY).  
FT DISULFID 236 239 BY SIMILARITY.

FT DISULFID 248 475 BY SIMILARITY.  
FT DISULFID 288 296 BY SIMILARITY.  
FT CARBOHYD 39 39 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 612 AA; 65486 MM; CDTB23B5FA978F97 CRC64;  
Query Match 69.8%; Score 1935; DB 1; Length 612;  
Best Local Similarity 70.3%; Pred. No. 1.9e-120;  
Matches 367; Conservative 59; Mismatches 90; Indels 6; Gaps 4;  
QY 1 MSFRLALSLGLVCTGLA--NYSKRALTLDSWLSNEATYARPAIILNIGADGAWYSGADS 58  
DB 2 VFSFSCLPALAGSSVLAVQPVLRQATGIDTLSTEAFNSRQAIILNIGADGQSAQGASP 61  
QY 59 GIVVASPSTNDYPTWTPTROSGVLKTLVDFRRGDPSTLSTINYSQAQIVGQISNP 118  
DB 62 GVVIASPSKSDPDYFTWTRDSGLVWKLVDFRRGDDADLPFIIEFFISSQRIQGISNP 121  
QY 119 SGDLSSGAGLGEPKENVDETAYTGSMGRPORDPALRATAMIGFCQMLDNGYSTATDI 178  
DB 122 SGALSSG-GLGEPKENVDETAYTGSMGRPORDPALRATAMISPEMLVENSHTSIATDL 180  
QY 179 VWPVLRNDLSYVAQYWNQTDVLMEEVNGSGFTTAVQHRALVEGSAFATVAGSSCQWD 238  
DB 181 WPPVVRNDLSYVAQYWSQSGFDLMEEVNGSTSEFTVAVSHRALVEGSSFAKTGSSCPCYCD 240  
QY 239 SQAPRILCYLGSFMTGSPILANFDSRSRQKANTLIGSHTEPPEACDSTFOPCSPPA 298  
DB 241 SQAPVRCYLGSEFMTGSPILANFDSRSRQKANTLIGSHTEPPEACDSTFOPCSPPA 300  
QY 299 LANHKEVVDSPFRSITLNDGLSDSEAVAVAGRPEDTYVNGNPFCTLLAAEQLYDALYQ 358  
DB 301 LANHKEVVDSPFRSITLNDGLSDSEAVAVAGRPEDTYVNGNPFCTLLAAEQLYDALYQ 360  
QY 359 WDKXGSLVETDYSLDFFPALYSDAATGYSSSSSTYSIYDAVKTTPADGFIIVETHAAS 418  
DB 361 WDKIGSLAITVYSLPFFPALYSNAATGYASTYKDIVSAVKAAYGAYQIVQTVYAA 420  
QY 419 NSGSEQYDKSPGEBLSARDLTWYSYALTLNNRNSVVPASWGTSAVSVCCTAAISA 478  
DB 421 TGSMAEQYTKTDGSGTSDRLTWSYALTLNNRNSVVPASWGTSAVSVCCTAAISA 480  
QY 479 IGTYSVVTWSPSIVANAGT--TTTATPTGSGVSTSTKTT 518  
DB 481 SGTYSVVTWSPSIVANAGT--TTTATPTGSGVSTSTKTT 521  
RESULT 5  
AMYG NEUCR STANDARD; PRT; 626 AA.  
AC P14804; QP9505; -  
DT 01-APR-1990 (Rel. 14, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Glucosylase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)  
DE (1,4-alpha-D-glucan glucanohydrolase).  
GN GLA-1 OR B5022.70.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetiales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A. / FGSC 987;  
RX STRAIN=74-OR23-1A / FGSC 987;  
RX MEDLINE=94037144; PubMed=8221928;  
RA Stone P.J., Makoff A.J., Parish J.H., Radford A.;  
RT "Cloning and sequence analysis of the glucosylase gene of Neurospora  
RT crassa";  
RL Curr. Gene. 24:205-211(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=74-OR23-1A / FGSC 987;  
RX MEDLINE=22542210; PubMed=12655011;  
RA Manhaupt G., Montreone C., Haase D., Mewes H.-W., Altmann V.,

RA Hoheisel J.D., Fartmann B., Nyakatura G., Kempen F., Maier J.,  
 RA Schulte U.;  
 RT "What's in the genome of a filamentous fungus? Analysis of the  
 RT Neurospora genome sequence.";  
 RT Nucleic Acids Res. 31:1944-1954(2003).  
 RN [3]  
 RP SEQUENCE OF 36-65.  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RA Koh-luar S.I., Parish J.H., Bleasby A.J., Pappin D.J.C., Ainley K.,  
 RA Johansen F.E., Radford A.J;  
 RT "Exported proteins of *Neurospora crassa*: 1-glucosylase.";  
 RL Enzyme Microb. Technol. 11:692-695(1989).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-  
 CC glucose residues successively from non-reducing ends of the chains  
 CC with release of beta-D-glucose.  
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL67291; CAA47707.1; -;  
 DR EMBL; AL355932; CAB91426.1; -;  
 DR HSSP; P04064; 1GAI.  
 DR InterPro; IPR002044; CBD\_4.  
 DR InterPro; IPR008291; Glu-a-glcad\_SBD.  
 DR InterPro; IPR001651; Glyco\_hydro\_15.  
 DR InterPro; IPR009281; Glyco\_trans\_6np.  
 DR Pfam; PF00686; CBM\_20; 1.  
 DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
 DR PIRSF; PIRSF01031; Glu-a-glcad\_SBD; 1.  
 DR PRINTS; PR00736; GLHYDRLAS15.  
 DR PRODOM; PD001568; CBD\_4; 1.  
 DR PROSITE; PS00820; GLUCOAMYLASE; 1.  
 KM Hydrolyase; Glycosidase; Polysaccharide degradation; glycoprotein;  
 KM Signal.  
 FT SIGNAL. 1 19 POTENTIAL.  
 FT PROPEP 20 35 POTENTIAL.  
 FT CHAIN 36 626 GLUCOAMYLASE.  
 FT BINDING 155 155 SUBSTRATE (BY SIMILARITY).  
 FT ACT\_SITE 211 211 CATALYTIC BASE (BY SIMILARITY).  
 FT ACT\_SITE 214 214 GENERAL ACID CATALYST (BY SIMILARITY).  
 FT ACT\_SITE 215 215 INTERACT WITH SUBSTRATE (BY SIMILARITY).  
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 82 82 MISSING (IN REF. 1).  
 FT CONFLICT 550 550 A -> R (IN REF. 1).  
 FT CONFLICT 560 560 V -> L (IN REF. 1).  
 SQ SEQUENCE 626 AA; 66474 MW; 54B5BDBA7A3E349 CRC64;

Query Match 56.9%; Score 1577; DB 1; Length 626;  
 Best Local Similarity 57.6%; Pred. No. 8.6e-97;  
 Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

DB 192 PIYKNDLAVYATQYNNNTGPFDMEEVANSSEFTTAAASHRALVEGSAFAKSSGSSACDAI 251  
 QY 241 APEILCYLGSPWGS-FILANFDSRSRGKANTLGSINTFPDPAACDSTFQPCSPRAL 299  
 DB 252 APOILCFQGSFWSNGYIIENFNVYRSGKDINSVLTIHNFDPACGDVNTFQCSDRAL 311  
 QY 300 ANHKVVDSPRSIYTLNDGSDSEAVAVERYPEDTYNGMPWLCTLAAEQLYDALYOM 359  
 DB 312 ANHKVVDSPSR-FMGVNSGRTAGAAAVGRADVYNGMPWLATLAAEQLDAYVW 370  
 QY 360 DKQGSLEVTNVLDPFALVSDAATGYSSSSSTYSISIVAVKTFADGFVSIETHAASN 419  
 DB 371 KKQGSITVTSIALFFDVLPSVSTGYSSSSSTYTHIINAVTTYADGFVDIYQYRPSD 430  
 QY 420 GSNSEQYDKSDGSGRLSARDLTWVSYALLTANNRNSVVPASWETSASVPGTCATSAI 479  
 DB 431 GSLAEQDKDSGAPLSATHTLTWSYASPLSAARRAGIIPPSWGAASANSIPGSCASTVA 490  
 QY 480 GTYSSTVTSWPSIVATGCTTT-ATPTGSGS-----VTSTKTTAASKT 524  
 DB 491 GSATATATSPFANLTTPASTVTPPTGTGCAADHEVLVTNEKVTYSYGT 541

RESULT 6  
 AMTG HORRE STANDARD; PRT; 616 AA.  
 ID AMTG HORRE  
 AC 003045;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Glucosylase P precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)  
 DE (1,4-alpha-D-glucan glucosylase).  
 GN GAMP.  
 OS Hormoonis resiniae (Creosote fungus) (Amorphotheca resiniae).  
 OC Eukaryote; Fungi; Ascomycota; Ascomycota incertae sedis;  
 OC Amorphothecaceae; Amorphotheca.  
 OX NCBI\_TaxID=5101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 20495;  
 RX MEDLINE=93146382; PubMed=1490604;  
 RA Joutsen J.V., Torkkeli T.K.;  
 RT "Glucosylase P gene of *Hormoonis resiniae*: molecular cloning,  
 RT sequencing and introduction into *Trichoderma reesei*.";  
 RL FEWS Microbiol. Lett. 78:237-243(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 20495;  
 RX MEDLINE=93365035; PubMed=8358830;  
 RA Vainio A.E.I., Torkkeli H.T., Tuusa T., Aho S.A., Fagerstrom B.R.,  
 RA Korhola M.P.;  
 RT "Cloning and expression of *Hormoonis resiniae* glucosylase P cDNA in  
 RT *Saccharomyces cerevisiae*.";  
 RL Curr. Genet. 24:38-44(1993).  
 RN [3]  
 RP SEQUENCE OF 72-76, AND CHARACTERIZATION.  
 RC MEDLINE=9038987; PubMed=2116499;  
 RA Fagerstrom R., Vainio A.E.I., Suoranta K., Pakula T., Kalkkinen N.,  
 RA Torkkeli H.T.;  
 RT "Comparison of two glucosylases from *Hormoonis resiniae*.";  
 RL J. Gen. Microbiol. 136:913-920(1990).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-  
 CC glucose residues successively from non-reducing ends of the chains  
 CC with release of beta-D-glucose.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>



DR PRINTS; PR00736; GLHYDLASE15.  
 KW Hydroxylase; Glycosidase; Polysaccharide degradation; Glycoprotein;  
 KM Signal; Meiosis.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT PROPER 17 28 BY SIMILARITY.  
 FT CHAIN 29 450 PROBABLE GLUCOMYLASE.  
 FT BINDING 147 147 SUBSTRATE (BY SIMILARITY).  
 FT ACT\_SITE 203 203 CATALYTIC BASE (BY SIMILARITY).  
 FT ACT\_SITE 206 206 GENERAL ACID CATALYST (BY SIMILARITY).  
 FT ACT\_SITE 207 207 INTERACT WITH SUBSTRATES (BY SIMILARITY).  
 FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 450 AA; 5163 MM; 31C5B2ABE4785FE CRC64;

Query Match 28.2%; Score 781.5; DB 1; Length 450;  
 Best Local Similarity 37.6%; Pred. No. 1.9e-44;  
 Matches 179; Conservative 68; Mismatches 166; Indels 63; Gaps 12;

QY 7 LALSGLVCTG---LANVSKRATIDSWLSNATYARLILNIGADGAWGAGSGIVVA 63  
 DB 8 LLLGGVVASBSLLSPNKRSKASMDWDQKGLAMGMLNINIGDSGHANDINPGCIIA 67  
 QY 64 SPSTDNDFYFTYTRDGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSGDL 123  
 DB 68 SPSTDSFDYYQWRDASLTITMTLDRFEEDKGLPIYKYMDEMRLOQVNPSPGDFY 127  
 QY 124 SGAGLSEPKFNDTAYTGWGRPQDGPALRATAMTGPQMLDNGYTSTATDIWPLV 183  
 DB 128 AG-GLGEPKFNVDGTSYDGMGRPQDOSPALRAIAFIKYNMYLPENGKEVEHVTWIEAV 186  
 QY 184 RNDLSYVAQYVNGQYGLMEVNGSSPFTIVQHRALVEGSAFPAVAGSSCKDSQAP 243  
 DB 187 LADLDYTAHMTWTFSPDLMEIKDVHYFTLAVQKRAMODGTAFAKRG-----APD 237  
 QY 244 -----ILCYLQSFV-TGSEFILANFDS--SRSGKDNLTLSGSIHT--PDPEAACD 287  
 DB 238 QALYQRTIEPIDKLGEFMDPGMGVIGYKGRDRSGLDSTLLASISYSEFD----- 291  
 QY 288 DSTFQPCSPRALANHKEVDSFNSIYTLNDGLSDSEAVAGRYEDTYN-----GNPWF 342  
 DB 292 -----MHLLPTLLKLTQETMTDRDYPVNGMKQ-----AMGRYEDVDYDGVSKSIGNPWF 339  
 QY 343 LCTLAAEQLYDALYQMDKQSLVTDVSLDFK--ALYSAAAGTYSSTSSSTYSYVDA 400  
 DB 340 ICTSSAAEIIYKALAYDNKGLPELTENYHFMKFAEFGD-----PYNMSYIRKN 390  
 QY 401 VKTPADGFVSIVETHAASNGMSSEQYDKSDGEQLSARDLTYSYALLTANRRRSV 456  
 DB 391 MHTYADNFLKAVAFQHPNGSMSEQFSRDGHOKGANDLTYSYSLNATYRRAI 446

RESULT 8  
 AMYG RHIOR STANDARD; PRT; 604 AA.  
 AC P07683;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Glucosylase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)  
 OS Rhizopus oryzae (Rhizopus dellemat).  
 OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;  
 OC Rhizopus.  
 OX NCBI\_TaxID=64495;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SAM0034;  
 RA Ashikari T., Nakamura N., Tanaka Y., Kiuchi N., Shibano Y.,  
 RA Tanaka T., Amachi T., Yoshizumi H.,  
 RT "Rhizopus raw-starch-degrading glucosylase: its cloning and  
 RT expression in yeast."  
 RL Agric. Biol. Chem. 50:957-964(1986).  
 RN [2]

RP HOMOLOGY, AND PREDICTED SECONDARY STRUCTURE.  
 RA Tanaka Y., Ashikari T., Nakamura N., Kiuchi N., Shibano Y.,  
 RA Amachi T., Yoshizumi H.,  
 RT "Comparison of amino acid sequences of three glucosylases and their  
 RT structure-function relationships."  
 RL Agric. Biol. Chem. 50:965-969(1986).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-  
 CC glucose residues successively from non-reducing ends of the chains  
 CC with release of beta-D-glucose.  
 CC -1- MISCELLANEOUS: Rhizopus glucosylase exists in multiple forms,  
 CC Gluc 1, Gluc 2, and Gluc 3, all of which hydrolyze gelatinized  
 CC starch at similar rates, but only the largest one (Gluc 1) is able  
 CC to adsorb raw starch.  
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL; D000049; BAA00033.1; -.  
 CC PIR; JP0001; JP0001.  
 DR HSSP; P04064; IGAT.  
 DR InterPro; IPR005036; CBM\_21.  
 DR InterPro; IPR00165; Glyco\_hydro\_15.  
 DR InterPro; IPR008928; Glyco\_trans\_6hp.  
 DR Pfam; PF03370; CBM\_21; 1.  
 DR Pfam; PF00723; Glyco\_hydro\_15; 1.  
 DR PRINTS; PR00736; GLHYDLASE15.  
 DR PROSITE; PS00820; GLUCOMYLASE; 1.  
 KW Hydroxylase; Glycosidase; Polysaccharide degradation; Glycoprotein;  
 KM Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 26 604 GLUCOMYLASE 1 (GLUC 1).  
 FT CHAIN 159 604 GLUCOMYLASE 2 (GLUC 2).  
 FT CHAIN 116 604 GLUCOMYLASE 3 (GLUC 3) (OR 110-604).  
 FT DOMAIN 26 109 ADSORPTION TO RAW STARCH (OR 26-115).  
 FT DOMAIN 116 604 STARCH DEGRADATION (OR 110-604).  
 FT BINDING 279 279 SUBSTRATE (BY SIMILARITY).  
 FT ACT\_SITE 336 336 CATALYTIC BASE (BY SIMILARITY).  
 FT ACT\_SITE 339 339 GENERAL ACID CATALYST (BY SIMILARITY).  
 FT ACT\_SITE 340 340 INTERACT WITH SUBSTRATES (BY SIMILARITY).  
 SQ SEQUENCE 604 AA; 65162 MM; 78421FLAA3ADB9 CRC64;

Query Match 26.2%; Score 725; DB 1; Length 604;  
 Best Local Similarity 36.8%; Pred. No. 1.6e-40;  
 Matches 165; Conservative 71; Mismatches 170; Indels 42; Gaps 10;

QY 25 ATLDSWLSNEATYARLILNIGADGAWGAGSGIVVASTDNPDPYFTWTRDGLVL 84  
 DB 168 STISSWIKKEGISRPMMLNINP-----FGSATGFRAALSTGPRPYTAMTRDALTS 222  
 QY 85 KTLVDLFR---NGDTSLSTIENYISAQAIVQGISNPSGDLSEAGLSEPKFNDTAYT 141  
 DB 223 NVIYEYNTLTSGKTLTANLKDVTFPSVKTQSTSTVCN-----CLGSEKFNPDAGYT 276  
 QY 142 GSWGRPRDRPALRATAMTIFG--QMLDNGYTSTATDIWPLRNDLSYVAQYVNGQYD 200  
 DB 277 GAWGRPQNDPAPAEATYFLLFADSYLTQTRDASVYGTLLPAFKDLDYVNVWNSGCFD 336  
 QY 201 LMEEVNGSPFTTIAVQHRALVEGSAFPAVAGSS--CSMCDSQAPELICYGSFWTGSFIL 258  
 DB 337 LMEEVNGVHYTTILMWRKGLIGADPAKRNGDSTRASTYSTASTANKISSFWSSNMW 396  
 QY 259 ANPDS-----SRSGKDANTL---LGSIHFPDEAACDSTFQPCSPRALANHKEVDS 308  
 DB 397 IYVSQSVTGAVSKKGLDVSTLLANLGSV-----DDGFFPGSEKILATVAVEDS 447  
 QY 309 FRSLYTLNDGLSSEAVAVGRYPEDY-----YNGNPFELCTLAAEQLYDALYQMDKQ 363

DB 448 FASTVPINKNLPYSIGNSIGRYPEDTYNGNSQSGNFWFLATVGAELYRAIKEWIGNG 507

QY 364 SLEVPDVSLEDFKALYSDAATG-TYSSSSSTYSIVDAVKTADGVSVIYETHAASNGM 422

DB 508 GTVSSISLPPFKKDDSSATSGKTYVTGSDNNLAQNIALADRLSTVQJAHANNGL 567

QY 423 SEQYKSDQEQSLARDLTWVSAAALTAN 450

DB 568 AEFPRITGGLSTGARDLTWVSAAALTAN 595

RESULT 9

AMYH\_SACFI ID AMYH\_SACFI STANDARD: PRT: 519 AA.

AC P2698: P78745; 15-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Glucoamylase GLA1 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).

CN GLA1.

OS Saccharomycopsis fibuligera (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.

OX NCBI\_TaxId=4944;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=KZ;

RA MEDLINE=92137640; PubMed=1840532;

RA Hostinova E., Balanova J., Gasperik J.;

RT "The nucleotide sequence of the glucoamylase gene GLA1 from Saccharomycopsis fibuligera KZ."

RL PEMS Microbiol. Lett. 67:103-108(1991).

RN [2]

RP REVISIONS.

RC STRAIN=KZ;

RA Hostinova E.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose.

CC -1- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES. ONLY DIFFERS IN 11 POSITIONS WITH GLUCOAMYLASE GLU1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X58117; CAA41120.1; -.

DR HSSP; P08017; IAYX.

DR InterPro; IPR000165; Glyco\_hydro\_15.

DR InterPro; IPR008928; Glyco\_trans\_15.

DR Pfam; PF00723; Glyco\_hydro\_15; 1.

DR PRINTS; PR00736; GLHYDRLASE15.

DR PROSITE; PS00820; GLUCOAMYLASE.1.

KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein; Signal.

FT SIGNAL 1 27

FT CHAIN 28 519

FT BINDING 166 166

FT ACT\_SITE 234 234

FT ACT\_SITE 237 237

FT ACT\_SITE 238 238

FT CARBOHYD 115 115

FT CARBOHYD 127 127

FT CARBOHYD 205 205

SO SEQUENCE 519 AA; 57542 MW; 1A5A009A7640053C CRC64;

Query Match 24.2%; Score 671.5; DB 1; Length 519;

Best Local Similarity 35.2%; Pred. No. 4,3e-37;

Matches 173; Conservative 73; Mismatches 169; Indels 77; Gaps 16;

QY 24 RATLDSWLSNEATVARTALINNIGADGAVSGADSGIVVASSTNDPDIYETWTRDGL- 82

DB 42 RDLTETFLDKQDVSLVYLQNIAYPEQGFNDGVGVGLVASSPSTNPDIYQMTDRDAIT 101

QY 63 ---VKTIVDLPRNGDLSLTENIYISAGALVQGISNPSG--DLSAGAGLGPENNVDE 137

DB 102 FLTVLSELD--NNFNTLAKAVEYIINTSYNLQRTSNPSGSFDDENHGLGEPKFNITG 159

QY 138 TAYTSGMRPQRDGPALRATAM-----IGFQMLL-----DNGYTSATDIWPLVR 184

DB 160 SAYTAMGRPQNDGALAAAYALSRNLVDNLSLANKLVLTGSDGDNFST-EDIYKNIIL 218

QY 185 NDLSTVAOYMNQGYDLMEEVNGSSFFTIYVQRLV-----EGSAFPAVAGSSCS 235

DB 219 PLEEVYIGWSTGDTLWENQGRHFFTSVQKLAAYADIAKSFDDGDFANTLSSTAS 278

QY 226 WDSQAPRLICLOSFWTGS-----FLANFP-----SSRGKANTLIGSIHFPD 281

DB 279 -----TLSEYLSGSDGFPVNTDVNHLVENPDLQONSQGLDSATYIGPLTHD 327

QY 282 PEACDDSTPQCPSPRALANHKVEYVDSFRSIYTLNDGLSDSPAVALGRPEPTY-----Y 336

DB 328 -IGESSSTPFVDNNTVYLOSTYLLIEDNKRYSVNSAY--\$AGAAIGRYPEVYNGDGS 384

QY 337 NGNPWFCTLAAAEQLYDALYQMDKQSLLEV--DVSLDFKALYSDAAT----- 384

DB 385 EBNPWFFLATVAAQVPYKLVVD--ASASNDITINKINVDFFNKIYVDLSTINGYQSSDS 443

QY 385 GYSSSSSTYSIVDAVKTADGVSVIYETHAASNGSEQYKSDQEQSLARDLTWVS 444

DB 444 VTIKSGDSFNTVAANLVTFGDSFLQVLDHINDGSINBOLNRTGYSTSYSLTWSSG 503

QY 445 ALLTANRRNSY 456

DB 504 ALLTAIRLRNKY 515

RESULT 10

AMYG\_SACFI ID AMYG\_SACFI STANDARD: PRT: 519 AA.

AC P08017;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glucoamylase Glu1 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).

CN GLU1.

OS Saccharomycopsis fibuligera (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.

OX NCBI\_TaxId=4944;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=87307999; PubMed=3114236;

RA Itoh T., Ohtsuki I., Yamashita I., Fukui S.;

RT "Nucleotide sequence of the glucoamylase gene GLU1 in the yeast Saccharomycopsis fibuligera."

RL J. Bacteriol. 169:4171-4176(1987).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).

RC MEDLINE=98437615; PubMed=9757101;

RA Sevcik J., Solovicova A., Hostinova E., Gasperik J., Wilson K.S., Dauter Z.;

RT "Structure of glucoamylase from Saccharomycopsis fibuligera at 1.7-A resolution."

CC Acta Crystallogr. D 54:854-866(1998).

CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose.

CC -1- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES. ONLY

CC DIFFERS IN 11 POSITIONS WITH GLUCOAMYLASE GLA1.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
-----  
DR EMBL: L25641; AAA83997.1; -.  
DR EMBL: M17355; AAA34649.1; -.  
DR PDB: 1AYX; 13-MAY-98.  
DR InterPro: IPR00165; Glyco\_hydro\_15.  
DR InterPro: IPR008928; Glyco\_trans\_6hp.  
DR Pfam: PF00723; Glyco\_hydro\_15; 1.  
DR PRINTS: PR00736; GLHYDRLASE15.  
DR PROSITE: PS00820; GLUCOAMYLASE; 1.  
KM Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;  
KW Signal; 3D-structure.  
FT CHAIN 1 27  
FT BINDING 28 519 GLUCOAMYLASE GLU1.  
FT ACT\_SITE 166 166 SUBSTRATE (BY SIMILARITY).  
FT ACT\_SITE 234 234 CATALYTIC BASE (BY SIMILARITY).  
FT ACT\_SITE 237 237 GENERAL ACID CATALYST (BY SIMILARITY).  
FT ACT\_SITE 238 238 INTERACT WITH SUBSTRATES (BY SIMILARITY).  
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT TURN 30 31  
FT TURN 36 37  
FT TURN 42 43  
FT HELIX 45 61  
FT TURN 62 63  
FT STRAND 64 65  
FT TURN 67 68  
FT STRAND 71 71  
FT TURN 76 77  
FT STRAND 79 79  
FT STRAND 85 85  
FT STRAND 89 89  
FT STRAND 93 94  
FT HELIX 95 111  
FT TURN 112 113  
FT HELIX 116 133  
FT TURN 134 134  
FT STRAND 137 137  
FT TURN 138 139  
FT STRAND 140 140  
FT TURN 142 143  
FT HELIX 145 151  
FT STRAND 154 155  
FT TURN 157 158  
FT STRAND 161 161  
FT TURN 171 171  
FT HELIX 173 191  
FT TURN 193 194  
FT TURN 199 200  
FT TURN 202 203  
FT HELIX 209 215  
FT TURN 216 216  
FT HELIX 217 227  
FT TURN 228 229  
FT STRAND 232 233  
FT TURN 235 236  
FT STRAND 240 241  
FT HELIX 244 263  
FT TURN 264 265  
FT HELIX 267 285  
FT TURN 287 289  
FT STRAND 292 293  
FT TURN 294 297  
FT STRAND 298 299

FT HELIX 303 307  
FT TURN 308 309  
FT TURN 316 316  
FT HELIX 317 325  
FT TURN 328 329  
FT TURN 338 339  
FT HELIX 341 357  
FT TURN 359 363  
FT TURN 374 375  
FT STRAND 378 378  
FT STRAND 383 383  
FT STRAND 389 389  
FT HELIX 390 410  
FT TURN 411 411  
FT STRAND 414 417  
FT HELIX 418 420  
FT HELIX 421 427  
FT TURN 428 428  
FT HELIX 431 433  
FT TURN 434 434  
FT TURN 436 437  
FT TURN 438 440  
FT STRAND 443 446  
FT TURN 448 449  
FT HELIX 451 474  
FT TURN 477 478  
FT STRAND 484 485  
FT TURN 487 489  
FT STRAND 492 493  
FT TURN 496 497  
FT HELIX 499 519  
SQ SEQUENCE 519 AA; 57539 MW; BE73035AD1B77652 CRC64;  
  
Query Match 24.0%; Score 664.5; DB 1; Length 519;  
Best Local Similarity 34.8%; Pred. No. 1.3e-36;  
Matches 171; Conservative 73; Mismatches 171; Indels 77; Gaps 16;  
  
QY 24 RATLDWLSNEATVARTAIINIGADAWVSGADSGIVVASFPTDNPDEYFYTWRDGL- 82  
D 42 RTDLETFRLDKQKEVSLVYLLQNTIAYPEGQFNNGVPGTVIASPSTNSPDYIYQWTRDSALT 101  
QY 83 ---VLKTLVDLFRNGDTSLSITENYISQAIVQGISNBSG--DLSSGALIGERKFVDE 137  
D 102 FLTVLSELED--NNFNNTLAKAVEYUINTSYNLRQTSNPSGSPDENHKGIGEPKFTD 159  
QY 138 TAYTGSWGRPORDGPALRATAM-----IGRGWLL-----DNGYSTATDIPVLR 184  
D 160 SAYTGAMGRPQNDGPALRAVAISRYLNDVNSLNEGKVLTDGDLNFSST--EDLYKNIIK 218  
QY 185 NDLSYVAQVNNQGYDLMEEVNGSSPFTIAVQHRALV-----EGSAFATAVGSSES 235  
D 219 PDLEEVIGVDSGTGFDLMENQGRHFFTSLVQOKALAYAVDIKSPDDGDFANTLSSTAS 278  
QY 236 WCDSQAPELICYLOSEFTGS-----FILANFD-----SRSGCANATLGSITTFD 281  
D 279 -----TLESYSGSDGGFVNTDVNHIYENPDLLQONSROGSDSATYIGPLLTID 327  
QY 282 PBAACDSTFQPSRALANHKVVDSEFRSITYLLNDGLSDSEAVAVGRYEDPY-----Y 336  
D 328 -IGSSSTPDPVDNEVYLOSITYLLEDNKKRYSVNSAY--SAGAAIGRYEDDYNGDSS 384  
QY 337 NGNPWFELCTLAABEQLYDALYQWDKQGLEVT--DVSIDPF-----KALYSDPAT 384  
D 385 EGNPMFLATAYAAQVRYKALYD-AKASANDITINKINYDFNKIYVDLSTINSAYQSSDS 443  
QY 385 GYSSSSSTYSSIVDAVKTFADGFSIVETHAANSNGMSQYDKSDGEQLSARDLTWSYA 444  
D 444 VTIKSGSDEFNTVADNLVTFGDSFLQVYILDHINDGSLNQLNRYYTGYSTGAYSLTWSG 503  
QY 445 ALLTANNRRNSV 456  
D 504 ALLEAIRLRNKV 515





```

DR HSSP; P08017; 1AYX.
DR GerMOnline; 139634; -.
DR SGD; S0001361; SGAL.
DR GO; GO:0000324; C: vacuole (sensu Fungi); IDA.
DR GO; GO:0000439; F: glucan 1,4-alpha-glucosidase activity; IDA.
DR GO; GO:0007151; P: sporulation (sensu Saccharomycetes); IEP.
DR InterPro; IPR00165; Glyco_hydro_15.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR PRINTS; PR00736; GLHYDLASE15.
DR PROSITE; PS00820; GLUCAMYLASE; 1.
DR HydroLase; Glycosidase; Polysaccharide degradation; Sporulation.
KM BINDING 198 198 SUBSTRATE (BY SIMILARITY).
FT ACT_SITE 261 261 CATALYTIC BASE (BY SIMILARITY).
FT ACT_SITE 264 264 GENERAL ACID CATALYST (BY SIMILARITY).
FT ACT_SITE 265 265 INTERACT WITH SUBSTRATES (BY SIMILARITY).
FT CONFLICT 504 549 HYDROLYSIS OF TERMINAL 1,4-LINKED ALPHA-D-GLUCOSE RESIDUES SUCCESSIVELY FROM NON-REDUCING ENDS OF THE CHAINS WITH RELEASE OF BETA-D-GLUCOSE.
SQ SEQUENCE 549 AA; 61463 MW; 6351E84F2CFAAB77 CRC64;

Query Match 21.5%; Score 597; DB 1; Length 549;
Best Local Similarity 30.4%; Pred. No. 3.9e-32;
Matches 153; Conservative 89; Mismatches 202; Indels 60; Gaps 14;

QY 1 MSFSLALSGVCTGLANVISKATLDSWISNEATYATRALINIGADGAWSGADSGI 60
DB VQLRDVILAMNGTVYVD--SNGAMDSSALEWLOQKKVSIKIPENIGSAVYPS-ISP 113
QY 61 VVASPSNDPQFYTWTRDGLVKTIVDLFRNDTSLSTIENYISAGALVQGISNPSG 120
DB 114 VVASPSQTHDPYQWIRDSALTTNSIVS--HSAGPAIETLLQYLVNSFHLQSRNN-- 167
QY 121 DLSSAGC-----LGEPRFVDEATATYSGWRQQRDGPALRATAMIGQWMLDNG-- 170
DB 168 TLGAGIGTNDTVAGCPKKNVNDTATFEDMKGRQNDGPALRSALIKIIDYIQSGTDL 227
QY 171 -----YTSTATDIWPLVRNDLSVAYQWNGTGYDLMEVNGSSFFITAVQHRALVESA 225
DB 228 GAKYPFQSTA-DIPDIDIVRMDLRFIIDHWNSSGFDLMEVNGHFFTLVQLSAVDKSL 286
QY 226 FATVAGSSGSCWD---SQAPRILCYL---QSFPTGSFILANF-----DSSRSKDA 270
DB 287 YFNSSERSPPVEELRQTRDISKFLVDPPANGFINGRK--NYLVGTPMIADTLRSGLDI 343
QY 271 NTLLGSIHTFEDPEAACDSTFQPCSPALANHKHVVDSFRSITYLINDGLSDEAVAVARY 330
DB 344 STLLAANTVNDAPRA-SHLPDINDPRAVLNTHLMHLSIYIINDSSKRAIGALGRY 402
QY 331 PEDTY-----YNGNPWFLCTLAABEQLYDALYQMDKQSGLEVTDVSLDFFRALYSDAATG 385
DB 403 PEDYDYGFGEGEPWVLATCTASTLYQLYLRHISEQHDLVVPMNNDCSNAFMSSELVFS 462
QY 386 TYSS-----SSSTYSIVDAKTPADGFVSIVERTHAASNGSMSQYKSGOEG 433
DB 463 NLTTLGNDGYLLIEFTPAFNQTIQKIFOLDAPLVKLKAVGTGDELSEQFNKYTGFM 522
QY 434 LSARDLTWSYAAALLTANRRNSV 457
DB 523 QGAGHLLTWSYTFMDAYQIROEVL 546

RESULT 13
AMNH_SACDI STANDARD; PRT; 767 AA.
AC P0405; Q92314;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glucosamylase S1 precursor (EC 3.2.1.3) [Glucan 1,4-alpha-glucosidase]
DE (1,4-alpha-D-glucan glucohydrolase) (GAT).
GN STA1 OR DEX2 OR MAL5.
OS Saccharomycetes diastaticus (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

```

```

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=41870;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 60709;
RX MEDLINE=65104778; PubMed=3918017;
RA Yamashita I., Suzuki K., Fukui S.;
RT "Nucleotide sequence of the extracellular glucosamylase gene STA1 in the yeast Saccharomycetes diastaticus.";
RL J. Bacteriol. 161:567-573(1985).
RN [2]
RP SEQUENCE OF 1-54 FROM N.A.
RA Yamashita I., Suzuki K., Sakuzo F.;
RT "Proteolytic processing of glucosamylase in the yeast Saccharomycetes cerevisiae.";
RL Agric. Biol. Chem. 50:475-482(1986).
RN [3]
RP SEQUENCE OF 1-64 FROM N.A.
RA Shima H., Inui M., Akada R., Yamashita I.;
RT "Upstream regions of the yeast glucosamylase gene which are required for efficient transcription.";
RL Agric. Biol. Chem. 53:749-755(1989).
CC -! CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose.
CC -! SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL, outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.1sb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X02649; CA26487.1; ALT_INIT.
DR EMBL; D00428; BAA00332.1; -.
DR PIR; A21896; ALBYG.
DR HSSP; P08017; 1AYX.
DR InterPro; IPR00165; Glyco_hydro_15.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR PRINTS; PR00736; GLHYDLASE15.
DR PROSITE; PS00820; GLUCAMYLASE; 1.
KM HydroLase; Glycosidase; Polysaccharide degradation; Glycoprotein;
DR Signal; Multigene family.
FT CHAIN 1 21
FT DOMAIN 22 767
FT DOMAIN 22 347
FT DOMAIN 348 691
FT DOMAIN 692 767
FT BINDING 455 455
FT ACT_SITE 518 518
FT ACT_SITE 521 521
FT ACT_SITE 522 522
FT CARBOHYD 308 308
FT CARBOHYD 322 322
FT CARBOHYD 414 414
FT CARBOHYD 423 423
FT CARBOHYD 434 434
FT CARBOHYD 513 513
FT CARBOHYD 546 546
FT CARBOHYD 645 645
FT CARBOHYD 650 650
FT CARBOHYD 720 720
FT CARBOHYD 741 741
SQ SEQUENCE 767 AA; 82488 MW; A5F29E2427BDB593 CRC64;

Query Match 18.8%; Score 522; DB 1; Length 767;
Best Local Similarity 30.5%; Pred. No. 5.5e-27;
Matches 139; Conservative 80; Mismatches 177; Indels 60; Gaps 15;

```



OY 1 MSFSLALSGVCTGLANVSKRATLDSWLSNEATVARTALINNIGADGAMVSGADSGI 60  
 DB 313 VQLRVLVLMNGTVYD-SNGAMDSSPLEBWLOROKKVSIERIFENIGSAVYPS-ILRGV 370  
 OY 61 VVASPTDNDPYFYTWRTDGLVLTVDLFRNGDTSLSLTENYISAQAIYOGISNPSG 120  
 DB 371 VIASPSQTHPDYFYQWIRDSALTINSIVS--HSADPA-ITLLOQLNVSFHLQRTNN--- 424  
 OY 121 DLSGAG-----LGEPEKFNVDATYGSNGRQPRDGPALRATMIGCGMLDNG--- 170  
 DB 425 TLGAGIGTNDTVAGDPKMWNDVTAFTPEPGRPNPDGALSLILKIDIKOSGIDL 484  
 OY 171 -----YTSTATDIWPLVRNDSVYAQYWNQGYDLMEEVNCGSPFTIAYOHRALVEGSA 225  
 DB 485 GAKYFPGSTA-DIPDDIYRMDLRFIDHNSGDFLMEVNGMFFTLVQLSAVDRSL 543  
 OY 226 FATAVGSSGWCDC--SOAPEILCYL---OSFWTGSFIILNF-----DSRSRGDA 270  
 DB 544 YFNASERSSPFVEELRQTRDISKFLVDPANGFINGKY---NYIVETPMIADTLRSGLDI 600  
 OY 271 NTLGSIHTPPEAACDSTQPCSPRALANKEVDSFRSTYTINDGLSDEAVANGRY 330  
 DB 601 STLAAANTVHADPSA-SHLPEDINDPAVLTMLHMLNRSITYPIINDSSKNATGIALGRY 659  
 OY 331 PEDTY-----YNGNPFICTLAAAEQYDALYQMDKQSLSEVTVSLDFPKALYSDAATG 385  
 DB 660 PEDYVDGVEGEGNPMWLATCASTLYQLYRHISEQHDVVRANNDCSNAPSELYVPS 719  
 OY 386 TYSS-----SSSTYSSIVDAVKTFADGFV 409  
 DB 720 NLTTLGNDGYLLEFNTPAFNQTIQKIFQADSL 755

RESULT 14  
 AMYL\_SACDI STANDARD; PRT; 768 AA.  
 ID AC P29760;  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Glucoamylase S2 precursor (EC 3.2.1.3) (Glucon 1,4-alpha-glucosidase)  
 DB (1,4-alpha-D-glucan glucohydrolase) (GALT).  
 OS STR2 OR DEX1.  
 GN Saccharomyces diastaticus (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=41870;  
 RN RP SEQUENCE FROM N.A.  
 RX MEDLINE=94288586; PubMed=8017901;  
 RA Kim K., Bajzar G., Lee S.Y., Knudsen F., Mattoon J.R.;  
 RT "Cloning of a new allelic variant of a Saccharomyces diastaticus  
 RT glucoamylase gene and its introduction into industrial yeasts";  
 RL Appl. Biochem. Biotechnol. 44:161-185 (1994).  
 RN RP SEQUENCE FROM N.A.  
 RX MEDLINE=91276266; PubMed=2055484;  
 RA Landreche M.G., Pretorius I.S., Sollicet P., Marmur J.;  
 RT "Primary structure and regulation of a glucoamylase-encoding gene  
 RT Gene 100:95-103 (1991).  
 RL Gene 100:95-103 (1991).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-  
 CC glucose residues successively from non-reducing ends of the chains  
 CC with release of beta-D-glucose.  
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: M60650; AAA5107.1; -  
 DR EMBL: M90490; AAA20560.1; -  
 DR PIR: J00474; J00474.  
 DR HSSP: P08017; IAX.  
 DR InterPro: IPR000165; Glyco\_hydro\_15.  
 DR InterPro: IPR008928; Glyco\_trans\_6hp.  
 DR Pfam: PF00723; Glyco\_hydro\_15; 1.  
 DR PRINTS: PR00736; GHYDRAS15.  
 DR PROSITE: PS00820; GLUCOMYLASE; 1.  
 KW hydrolase; Glycosidase; Polysaccharide degradation; glycoprotein;  
 KW signal; Multigene family.  
 FT SIGNAL 1 21  
 FT CHAIN 22 768  
 FT DOMAIN 22 348  
 FT DOMAIN 349 692  
 FT BINDING 693 768  
 FT ACT\_SITE 456 456  
 FT ACT\_SITE 519 519  
 FT ACT\_SITE 522 522  
 FT ACT\_SITE 523 523  
 FT CARBOHYD 35 35  
 FT CARBOHYD 309 309  
 FT CARBOHYD 323 323  
 FT CARBOHYD 415 415  
 FT CARBOHYD 424 424  
 FT CARBOHYD 435 435  
 FT CARBOHYD 514 514  
 FT CARBOHYD 547 547  
 FT CARBOHYD 646 646  
 FT CARBOHYD 651 651  
 FT CARBOHYD 721 721  
 FT CARBOHYD 742 742  
 FT CONFLICT 164 164  
 FT CONFLICT 624 624  
 SQ SEQUENCE 768 AA; 82586 MW; 3FAC172C128A0C6F CRC64;  
 D -> N (IN REF. 1).

Query Match 18.8%; Score 521; DB 1; Length 768;

Best Local Similarity 30.5%; Pred. No. 6.4e-27;

Matches 139; Conservative 80; Mismatches 177; Indels 60; Gaps 15;

OY 1 MSFSLALSGVCTGLANVSKRATLDSWLSNEATVARTALINNIGADGAMVSGADSGI 60  
 DB 313 VQLRVLVLMNGTVYD-SNGAMDSSPLEBWLOROKKVSIERIFENIGSAVYPS-ILRGV 371  
 OY 61 VVASPTDNDPYFYTWRTDGLVLTVDLFRNGDTSLSLTENYISAQAIYOGISNPSG 120  
 DB 371 VIASPSQTHPDYFYQWIRDSALTINSIVS--HSADPA-ITLLOQLNVSFHLQRTNN--- 425  
 OY 121 DLSGAG-----LGEPEKFNVDATYGSNGRQPRDGPALRATMIGCGMLDNG--- 170  
 DB 425 TLGAGIGTNDTVAGDPKMWNDVTAFTPEPGRPNPDGALSLILKIDIKOSGIDL 485  
 OY 171 -----YTSTATDIWPLVRNDSVYAQYWNQGYDLMEEVNCGSPFTIAYOHRALVEGSA 225  
 DB 485 GAKYFPGSTA-DIPDDIYRMDLRFIDHNSGDFLMEVNGMFFTLVQLSAVDRSL 544  
 OY 226 FATAVGSSGWCDC--SOAPEILCYL---OSFWTGSFIILNF-----DSRSRGDA 270  
 DB 544 YFNASERSSPFVEELRQTRDISKFLVDPANGFINGKY---NYIVETPMIADTLRSGLDI 601  
 OY 271 NTLGSIHTPPEAACDSTQPCSPRALANKEVDSFRSTYTINDGLSDEAVANGRY 330  
 DB 601 STLAAANTVHADPSA-SHLPEDINDPAVLTMLHMLNRSITYPIINDSSKNATGIALGRY 660  
 OY 331 PEDTY-----YNGNPFICTLAAAEQYDALYQMDKQSLSEVTVSLDFPKALYSDAATG 385  
 DB 661 PEDYVDGVEGEGNPMWLATCASTLYQLYRHISEQHDVVRANNDCSNAPSELYVPS 720  
 OY 386 TYSS-----SSSTYSSIVDAVKTFADGFV 409  
 DB 721 NLTTLGNDGYLLEFNTPAFNQTIQKIFQADSL 756

```

RESULT 15
ID YG10 METVA STANDARD; PRT; 615 AA.
AC 059005;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical glycosyl hydrolase MJ1610 (EC 3.2.1.-).
GN MJ1610.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=868087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Rubmann J.L., Nguyen D.,
RA Utechtack T.R., Kelley J.W., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Colton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1. SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67601, AAB99630.1; -.
CC PIR: A64501, A64501.
CC TIGR: MJ1610; -.
CC InterPro: IPR000165; Glyco_hydro_15.
CC InterPro: IPR008928; Glyco_trans_6hp.
CC InterPro: IPR006465; Oligosac_amy1.
CC Pfam: PF00723; Glyco_hydro_15f.1.
CC TRGFAMS: TIGR01577; oligosac_amy1; 1.
CC PROSITE: PS00820; GUCCAMYLASE; 1.
CC Hypothetical protein; Hydrolase; Glycosidase; Complete proteome.
CC ACT_SITE 406 406 GENERAL ACID CATALYST (BY SIMILARITY).
CC FT ACT_SITE 407 407 INTERACT WITH SUBSTRATES (BY SIMILARITY).
CC SQ SEQUENCE 615 AA; 72007 MW; 2B37EB89F0357B85 CRC64;

Query Match 8.0%; Score 221.5; DB 1; Length 615;
Best Local Similarity 23.1%; Pred. No. 2.9e-07;
Matches 101; Conservative 46; Mismatches 125; Indels 165; Gaps 20;

QY 59 GIVVSPSTNDPDVEYMTTRSGVLKTLVDLPFRNGDTSLSLTENTYISAQIVQGISNP 118
DB 284 GGIILAPSL-HPDRIYMGKD-GSYISIALDLF-----GINNIPDRFEFEKSKIQ----- 331
QY 119 SGDLSSGAGIGEPKFNVDETAYTGSW-----GRPORDGPALRATAM---IGFGQWL 167
DB 332 -----NAD-----GSLQNYVYVNGK-----RLTALQTDQIGSIIMAM 364
QY 168 DNGYTSATDIWPLVVRNDLSYAQYMNQIG-----YDLMEEVNGSSFF 211
DB 365 DVHYRLT-----GDRKFEVERYVNTTEKAAHYLRVALNFTPCFDLMEERFPVAY 414
QY 212 TIANYQHALVEGSAFATAY-----GSSCSWCDSQAPRILCYLQSFWTGSIILANPD 262
DB 415 TMGATYAGLKCAYMSKAVKRRKRYVKMGKTIFFLKHEVPRKRP-YLR-----D 461

```

```

QY 263 SSRSGKDANTLLGSIHTPDEAACDSTFQPCSPRALANKEVYDSFRSITYTLNDGISDS 322
DB 462 EERFAKSIINPL-----DKTIDTSIIIGLSYFPNLIID-----VDDERMIXT-----A 501
QY 323 EAV-----AVGRYPEDTYNNGNPMFLCTLAAL---EQLYDALYQMDKQSLLEVTPV 370
DB 502 EALEKAFKRVGIGRIPEDIYFGGNPMWITTIMLSLYRRLKYVLEKDDNGA---DI 557
QY 371 SLDPFALYSDAATGYSSSSSTYSSIVDAVKTFADGFVSIVETHAASNGSMSEQYDKSD 430
DB 558 YLQSKKTLFWVWVKYSF-----DGLFPEQIHKEI 586
QY 431 GEQLSARDLTWSYALAI 447
DB 587 GVPMSAMPLGWSNAMFL 603

```

Search completed: June 17, 2004, 17:46:29  
Job time : 21 secs